

Inhalt

Vorwort	XI
---------	----

Speech, MMC

Angelika Salmen

Koordination multimodaler Metainformationen bei Fahrerinformationssystemen am Beispiel der Menüausgabe	1
---	---

Thomas Bauer, Ludwig Hitzenberger, Marcus Hennecke

Effects of Manual Phonetic Transcriptions on Recognition Accuracy of Streetnames	21
---	----

Informationsmanagement

Elisabeth Milchrahm

Entwicklung eines Modells zur Akzeptanzproblematik von Informationstechnologie	27
---	----

Christian Schlögl

Informations- und Informationstechnologie-Management in österreichischen Banken, Versicherungen und KFZ-Unternehmen	45
--	----

Mobile & Ubiquitous Information Access

Wolfgang Semar

Information überall, jederzeit	71
--------------------------------------	----

Maurice Danaher, Lennon Hopkins

Object Segmentation Using Colour for Security Applications on Mobile Devices	79
---	----

Lerntheorie & Wissensmanagement

Elisabeth Kamentz, Monika Schudnagis

Lerntheorie und kultureller Hintergrund - Einflussfaktoren bei der Gestaltung von Lernsystemen	85
---	----

<i>Stefan Berger, Franz Lehner</i> U-Know - Ubiquitäres Wissensmanagement an der Universität Regensburg	107
---	-----

<i>Gerhard Heyer, Thomas Hoppe, Katrin Müller</i> Die Wortschatz-Methode als netzbasierte Kreativitätstechnik in einer virtuellen Lernumgebung	123
--	-----

Visualisierung

<i>Maximilian Eibl, Markus Quandt</i> Modellbildung und Visualisierung: Das Sowinet.de-Wahlmodell .	131
--	-----

<i>Tobias Limbach, Frank Müller, Peter Klein, Harald Reiterer, Maximilian Eibl</i> Visualization of Metadata using the SuperTable + Scatterplot	147
--	-----

IR – Indexierung

<i>Lukas C. Faulstich, Uwe Quasthoff, Fabian Schmidt, Christian Wolff</i> Concept Extractor - Ein flexibler und domänenspezifischer Web Service zur Beschlagwortung von Texten	165
--	-----

<i>Libo Chen, Ulrich Thiel, Marcello L'Abbate</i> Automatische Thesauruserstellung und Query Expansion in einer E-Commerce-Anwendung	181
--	-----

IR, Web & Evaluation

<i>Joachim Griesbaum, Marc Rittberger, Bernard Bekavac</i> Deutsche Suchmaschinen im Vergleich: AltaVista.de, Fireball.de, Google.de und Lycos.de	201
---	-----

<i>Michael Kluck</i> Das Cross-Language Evaluation Forum (CLEF) - Evaluationsum- gebung und Forschungskontext für mehrsprachiges Information Retrieval (mit einer Skizze der Ergebnisse von CLEF 2002)	225
---	-----

<i>Thomas Mandl</i> Evaluierung von Internet-Verzeichnisdiensten mit Methoden des Web-Mining	239
IR & Fachinformation	
<i>Robert Strötgen</i> Behandlung semantischer Heterogenität durch Metadatenextraktion und Anfragetransfers	259
<i>Bettina Mielke</i> Wider gängige Ansichten zur juristischen Informationserschließung	273
Kooperation & Personalisierung in digitalen Bibliotheken	
<i>Peter Hanke, Thomas Mandl, Christa Womser-Hacker</i> Ein „Virtuelles Bibliotheksregal“ für die Informationswissen- schaft als Anwendungsfall semantischer Heterogenität	289
<i>Bernard Bekavac, Joachim Griesbaum, Rainer Kuhlen, Andreas Lenich, Thomas Schütz, Wolfgang Semar</i> ENFORUM - Stand und Aussichten	303
<i>Kurt Englmeier, Javier Pereira</i> Personal Assistance: Helping the Users Find Their Way through the Information Space	323
Web-Based Learning	
<i>Klaus J. Schäfer</i> Virtuelle Universität Regensburg Version 2 - Internet-Plattform für Learning on Demand	335
<i>Folker Caroli, Christa Hauenschild, Daniela Wagner</i> Netzgestützte Lehre - Vom Projekt zum Regellehrbetrieb	349

Demonstrationen & Poster

<i>Walter Piechulla, Christoph Mayser, Helmar Gehrke, Winfried König</i>	
Echtzeit-Fahrerbeanspruchungsschätzung	365
<i>Uwe Quasthoff, Matthias Richter, Christian Wolff</i>	
„Wörter des Tages“ - Tagesaktuelle wissensbasierte Analyse und Visualisierung von Zeitungen und Newsdiensten	369
<i>Steffen H. Elsner</i>	
Mehr oder weniger Mobilität durch Virtualität? Wissens- und Informationsmanagement über interaktive Netzwerke	373
<i>Maurice Danaher, Duy Nguyen</i>	
Mobile Home Security with GPRS	377
<i>Maurice Danaher, Chi Wai Lee</i>	
Creating a 3D University website with VRML	381
<i>Thomas Mandl, Maximilian Eibl</i>	
Ein prototypisches Werkzeug zur flexiblen Visualisierung komplexer Datenbestände	385
<i>Monika Schudnagis, Christa Womser-Hacker</i>	
SELIM: Softwareergonomie für Lernsysteme mit verschiedenartiger lerntheoretischer Basis	387
<i>A. Jackson, M. Lewandowski, Gr. Thurmair, J. Zwickl</i>	
ConceptManager: Pflege multilingualer Ontologien im crosslingualen Retrieval	389
<i>Ralph Kölle, Thomas Mandl, Christa Womser-Hacker</i>	
Implementierung und Evaluierung des MIMOR-Modells: Mehrfachindexierung zur dynamischen Methoden-Objekt- Relationierung im Information Retrieval	391
<i>Bernd Hermes, Maximilian Stempfhuber</i>	
MURBANDY - Benutzerfreundliche Recherche von Landnutzungsdaten	393
<i>Elke Lang, Andreas Bohne-Lang, Claus-Wilhelm von der Lieth, Alexander Loß</i>	
LINUXS - Eine Notation zur Unterstützung von Repräsentation und Retrieval spezieller chemischer Strukturen	395

Die modernen Industrie- und Informationsgesellschaften kennzeichnet eine permanent ansteigende Anforderungen an die Mobilität ihrer Mitglieder. Die daraus resultierenden Herausforderungen sind nur mit Hilfe eines optimierten Einsatzes von Informationssystemen zu lösen. So steht einerseits im Fokus die Optimierung der Mobilität bspw. von Fahrerassistenzsystemen, die durch Routenplanung oder Stauhinweise die Streckenführung verbessern sollen, während andererseits elektronische Kommunikation unnötige Mobilität vermeiden hilft. Hier sind beispielhaft Entwicklung und Bewertung netzbasierter virtueller Lehr- und Lernsysteme zu erwähnen. Diese und verwandte Themen werden in diesem Band behandelt.

Dieser Tagungsband enthält die wissenschaftlichen Beiträge des 8. Internationalen Symposiums für Informationswissenschaft »Information und Mobilität«.

(Regensburg, 8.-10. Oktober 2002)



Koordination multimodaler Metainformationen bei Fahrerinformationssystemen am Beispiel der Menüausgabe

Angelika Salmen

Universität Regensburg
Institut für Medien-, Informations- und Kulturwissenschaften
Universitätsstraße 31
93040 Regensburg
angelika.salmen@sprachlit.uni-regensburg.de

Zusammenfassung

Formen der multimodalen Menüausgabe wurden am Beispiel einer elektronischen Bedienungsanleitung hinsichtlich einer effizienten Systeminteraktion und der Verkehrssicherheit in einem realen Fahrversuch untersucht. Über das Systemdesign sollte die Nutzung der sprachlichen Interaktion als primärer Modus suggeriert werden. Als relevant erwies sich zunächst, dass die visuelle Menüausgabe konsekutiv zur sprachlichen erfolgt. Die Metainformationen sind soweit als möglich auf den sprachlich-akustischen Modus zu beschränken und auf die nötigsten Informationen zu reduzieren. Andernfalls führt dies zur Ablenkung der Benutzer und zu negativen Einflüssen auf das Fahrverhalten. Dennoch behält die Vermittlung von aktiven Systemzuständen für jede Komponente Relevanz.

Abstract

A variety of multimodal menu outputs were examined in a real driving situation with respect to efficient system interaction and aspects of safety. An electronic manual served as an exemplary application. The system has been designed with the prime intention of supporting the driver in interacting orally with the system. It proved expedient to display the menu after the spoken output had ended. Metainformation may possibly need to be presented in the acoustic mode and the amount of information required kept to a minimum as otherwise drivers become distracted and their driving behavior is negatively impacted. Conveying the actual states of the system is still relevant for each individual component.



1 Einleitung

In der Verkehrsinteraktion verläuft der Großteil der Informationsverarbeitung über den visuellen Kanal. Der akustische Kanal ist dagegen weitgehend ungenutzt. Daher bietet es sich an, Sprache als Bedienmodus zu nutzen. Sowohl Hören als auch Sprechen produzieren keine Interferenzen mit der visuellen Wahrnehmung oder mit motorischen Reaktionen (cf. FÄRBER 1987; BECKER ET AL. 1995; BENGLER 1995; FRANZ ET AL. 1992), wodurch sich die sprachliche Bedienung von Fahrzeugsystemen im Hinblick auf die Verkehrsinteraktion prinzipiell als überlegen erwies. Problematisch sind dennoch umfangreiche Informationsausgaben, wie z.B. längere Menüs. Zur Selektion eines Menüpunktes müssen die Optionen kurzzeitig memoriert werden, wobei die Memorierungskapazität jedoch stark limitiert ist. Bei verbalen Informationen erstreckt sich diese lediglich auf 5 ± 2 Einheiten (MILLER 1956; SIMON 1974). Der Auswahlprozess bei längeren Menüs erfordert generell eine hohe Konzentration und ist mental sehr belastend (SALMEN ET AL. 1999). Muss der Auswahlprozess zudem in die Verkehrsinteraktion integriert werden, kann die Koordination der Handlungen im Bereich der zentral mentalen Verarbeitung zu Überlastung und Reaktionsblockaden führen (cf. PASHLER & JOHNSTON 1998).

Zur Reduktion dieses Gefahrenpotentials kann eine zusätzliche visuelle Menüausgabe als Entlastung der mentalen Beanspruchung genutzt werden. Um die damit einhergehende visuelle Ablenkung einzugrenzen, muss bei den Fahrern jedoch das Bewusstsein vorhanden sein, die visuelle Komponente nur bei Bedarf und in angemessenen Verkehrssituationen zu nutzen. Daher sollte bereits über das Systemdesign ein entsprechendes Benutzerverhalten suggeriert werden.

Es ist anzunehmen, dass der Modus der Menüausgabe selbst zunächst nur über den akustischen Kanal erfolgen sollte, um die sprachliche Interaktion als Hauptmodus zu unterstützen. Eine anschließende visuelle Menüausgabe auf dem Display dürfte das erwünschte Benutzerverhalten hervorrufen, da die Displayausgabe nur Relevanz besitzt, sofern die Auswahl im akustischen Modus nicht vorgenommen werden konnte. Auch für die Metainformationen kann angenommen werden, dass deren Ausgabe über den akustischen Kanal generell den sprachlichen Interaktionsmodus unterstützen. Dennoch ist zu klären, welche Art von Metainformationen für welchen Modus relevant ist.

2 Die sprachliche Komponente

Da die Sprachausgabe als alleiniger Modus genügen soll, ist die Systeminteraktion zunächst auf den Modus der mündlichen Kommunikation auszurichten. Im Gegensatz zum visuellen Modus bestehen im akustischen Modus keine persistenten Anzeigen und Zeigemechanismen, es muss alles verbalisiert werden. Da dies sowohl zeitintensiv als auch ablenkend ist, ist die Information auf das Nötigste zu begrenzen.

Pulldown-Menüs als typische Vertreter einer graphischen Listenauswahl weisen eine umfangreiche Funktionalität auf. Neben der Präsentation der Auswahloptionen, die auch durch Icons oder Shortcuts ergänzt werden können, bestehen diese z.B. aus der Anzeige des Menükopfes, Indikatoren von Subdialogen und Untermenüs, die durch Punkte bzw. Pfeile realisiert sind, sowie aus Separatoren, die inhaltliche Zusammenhänge kennzeichnen. Welche Informationen sind nun für den akustischen Modus relevant?

Aus Gründen der Selbstbeschreibungsfähigkeit von Dialogen ist es erforderlich, dem Benutzer ein *Feedback* darüber zu geben, ob der Auswahlprozess auch zu dem gewünschten Ziel führt. Vor dem Hintergrund der noch nicht vollkommen ausgereiften Spracherkennung kann der Benutzer ferner sofort erkennen, ob seine Eingabe richtig oder falsch erkannt wurde.

Einen Faktor der Verunsicherung stellt das fehlende Wissen über die zu erwartende *Informationsmenge* dar, wie sich in den Benutzertests des Projekts SIMBA¹ zeigte. Hier wurde der Prototyp einer interaktiven Bedienungsanleitung (IABA), die auf reiner Spracheingabe und -ausgabe basierte, entwickelt. Die fehlende Angabe führte dazu, dass die Aufmerksamkeit der Probanden je nach individueller Toleranz zwischen 3 und 5 Menüpunkten aufrecht erhalten blieb. War das Menü länger, zeigte sich Ungeduld, und die Aufmerksamkeit sank. Ist der Umfang dagegen vorab bekannt, können die Ressourcen und die Memorierungsstrategie entsprechend eingeteilt werden, was gerade bei der Anwendung in der Verkehrssituation von enormer Bedeutung ist (LUDWIG 1998; SALMEN & TAUSCHEK 1999). Während bei graphischen Menüs sowohl explizite (Anzahl, Seitenmenge) wie implizite Optionen (Wahrnehmung des gesamten Menüs, Größe des Scrollbar-Buttons) existieren, den Menüumfang zu vermitteln, steht im akustischen Modus nur die Verbalisierung dieser Information zur Verfügung.

¹ Das Projekt SIMBA (Simulation des intelligenten maschinellen Beifahrers) wurde an der Universität Regensburg in Kooperation mit DaimlerChrysler von 1996 bis 1999 durchgeführt.

Graphische Auswahllisten ermöglichen es ferner, die *aktuelle Position* zu vermitteln. Ist das komplette Menü ersichtlich, erschließt sich die Listenposition in Relation zum Menüanfang und zum Menüende. Bei umfangreicheren Auswahllisten kann auf Hilfsmittel zurückgegriffen werden, wie etwa Seitennumerierungen oder Scrollbars, die durch den mitwandernden Button die relative Position anzeigen. Zwar existieren auch akustische Pendants zur Scrollbar, allerdings setzen diese i.d.R. eine Lernphase voraus (BREWSTER ET AL. 1994), was dem Anspruch der intuitiven Verständlichkeit, die gerade in der Verkehrssituation relevant ist, nicht entspricht. Geeignet für den akustischen Modus ist dagegen die Numerierung der Menüpunkte, die im Zusammenhang mit der Nennung des Menüumfangs auf die noch zu erwartende Informationsmenge schließen lässt.

Eine explizite Kommentierung der Menüpunkte, ob diese direkt zu einer Informationsausgabe führen oder zu einem *Untermenü* ist dagegen wenig probat, wobei auch der Verzicht auf diese Informationen lediglich als suboptimal gelten kann. Andererseits kommt der sprachlichen Menüausgabe die zusätzliche Funktion zu, die Inhaltsstruktur des Systems zu vermitteln. D.h. Benutzer werden damit rechnen, dass sich die Menüauswahl über mehrere Ebenen erstrecken wird. Orientiert man sich bei der Menüstrukturierung zudem an realen Objekten kann dadurch die Menüfolge und -tiefe implizit ausgedrückt werden: Bei IABA bestand die erste Ebene aus funktionalen Kategorien wie *Fahren* oder *Raumklima*. Auf der zweiten Ebene waren die zugehörigen Systeme aufgelistet, wie *Tempomat* oder *Klimaanlage*. Handelte es sich um komplexe Systeme, führte deren Aufruf zu einem weiteren Menü der einzelnen Funktionen. Bei einfachen Systemen wie etwa der *Sitzheizung* führte die Auswahl dagegen direkt zur Ausgabe des Erklärungstexts (SALMEN & TAUSCHEK 1999; SALMEN ET AL. 1999).

Weiterhin sollte das *Ende des Menüs* signalisiert werden. Zwar impliziert die Angabe des Menüumfangs und die Numerierung diese Information, es erfordert jedoch eine Transferleistung des Benutzers, was im Sinne der Aufgabenangemessenheit vermieden werden sollte. Aufgrund der Aufgabenangemessenheit sollte dem Benutzer auch der *Zeitpunkt der Eingabemöglichkeit* vermittelt werden. Da diese beiden Dialogschritte in direkter zeitlicher Abfolge stehen, können sie innerhalb einer Systemrückmeldung zusammengefasst werden. Ein akustisches Signal hätte den Vorteil, dass es kurz ist; allerdings ist dessen Semantik erst zu erlernen. Besteht Ungewissheit über die Bedeutung, was bei erstmaligen oder seltenen Benutzern der Fall sein kann, kann der Memorierungsprozess beeinträchtigt werden (BALLAS & MULLINS 1991).

Aus Gründen der Transparenz ist deshalb eine verbale Rückmeldung vorzuziehen.

Ein wesentliches Kriterium ist ferner die Vermittlung des Auswahlprozesses. Insbesondere im Bereich der Telephonie wird dies oft durch ausführliche verbale Instruktionen vorgenommen, die ihren Ursprung in der DTMF²-Technologie haben. Instruktionen der Art *Wenn Sie Hilfe brauchen, sagen Sie Hilfe* sind nicht nur redundant sondern auch ineffektiv (BALENTINE 1999). Da die Menüausgabe die Aktion der Auswahl impliziert, ist es ausreichend, eine einmalige Aufforderung am Ende des Menüs zu geben (LUDWIG ET AL. 1996).

Bei der Verbalisierung von Metainformationen ist generell darauf zu achten, dass diese so knapp wie möglich gehalten sind. Gerade bei häufigerem Gebrauch führen immer wiederkehrende wortreiche Anweisungen schnell zu Langeweile (DE VRIES & JOHNSON 1997) oder gar Enerviertheit, wenn Formulierungen über die nötigen Informationen hinausgingen (LUDWIG 1998). Nicht zuletzt spielt der Zeitfaktor eine entscheidende Rolle. Im Rahmen von IABA hat sich gezeigt, dass die in Tabelle 1 angeführten Feedbacks als minimale Information ausreichend und gleichzeitig unmissverständlich waren (SALMEN ET AL. 1999).

Feedback-Funktionen	IABA-Feedbacks
Erkannte Eingabe	<i>Thema</i>
Menüumfang	Es gibt <i>n</i> Themen:
Menüposition	[Numerierung der Themen]
Menüende & Eingabeaufforderung	Welches Thema wollen Sie?

Tab. 1: Sprachliche Metainformationen

3 Die visuelle Komponente

Die Motivation der visuellen Ergänzung liegt darin, die Memorierung der Menüpunkte zu unterstützen. Insofern stellt sich die Frage, ob die Wiederholung der Metainformationen im visuellen Modus nötig ist. Dagegen spricht das Prinzip, redundante und irrelevante Informationen zu vermeiden, da sie von den zentralen Punkten ablenken und unter Umständen die Informationsverarbeitung behindern können (MULLET & SANO 1995: 126-29; HASEBROOK 1995: 64-66). Andererseits sollte im Sinne der Selbstbeschreibungsfähigkeit

² DTMF steht für Dual Tone Multiple Frequency.

und der Konsistenz der Systemzustand auch über das Display jederzeit ersichtlich sein.

Das *Anfragefeedback* dient als Rückmeldung über die erkannte Eingabe und die darauf folgende Aktion. Bei korrekter Erkennung wäre die visuelle Wiederholung redundant, und bei einer Fehlerkennung wäre die Displayausgabe ohne Relevanz, wodurch darauf verzichtet werden kann. Als sinnvoll kann dagegen die Anzeige des Kontexts im Sinne einer Überschrift erachtet werden. Gerade in der Verkehrssituation können plötzlich auftretende Ereignisse die volle Aufmerksamkeit fordern, wobei über das Anfragefeedback der Kontext der abgebrochenen Menüauswahl direkt wieder hergestellt werden kann. Nicht unproblematisch ist die Darstellung jedoch bei mehrseitigen Menüausgaben. Da die Leseperformanz mit zunehmender Informationsmenge sinkt und die Information möglichst mit einem bzw. wenigen Blicken wahrgenommen werden soll, gilt eine maximale Anzahl von vier bis fünf Zeilen bzw. Menüpunkten als optimal (KOZMA 1991). D.h. alle Menüs, die länger sind, sollten als Teillisten auf mehrere Seiten verteilt werden. Aus Gründen der Konsistenz müsste die Kontextinformation auf jeder Seite erscheinen, was teilweise zu einer unnötigen Erhöhung der Seitenanzahl und auch zu einer Vermehrung der Informationsmenge führen würde. Da dies einer schnellen Informationsaufnahme entgegensteht, sollte auf die visuelle Darstellung verzichtet werden.

Die Angabe über die *Menülänge* ist ebenfalls als redundant und damit als verzichtbar zu betrachten. Der einmalige Verweis darauf sollte genügen. Der primäre Zweck im sprachlichen Modus ist ferner, dass der Benutzer seine Aufmerksamkeit entsprechend der zu erwartenden Informationsmenge einteilen kann. Die Persistenz der visuellen Ausgabe erfordert dagegen keine kontinuierliche Aufmerksamkeit. Relevant ist allerdings, die *Numerierung* der Menüpunkte beizubehalten. Insbesondere, wenn die Nennung der Nummer als alternative Eingabemöglichkeit dient, muss die Zuordnung zu den Menüpunkten explizit sein.

Erforderlich ist ferner die Signalisierung des *Menüendes* bzw. bei der Ausgabe von Teilmenüs die Anzeige der *Fortsetzung*. Standardisierte Darstellungsformen bei graphischen Benutzeroberflächen sind Navigationsbuttons, die mit Richtungspfeilen gekennzeichnet sind, wobei diese allein stehen oder in einer Scrollbar integriert sein können. Bei letzterer kann zudem die relative Position über den Positionsbutton abgelesen werden. Alternativ können die Seitenzahlen explizit ausgedrückt werden: entweder in der Form 1 v. 3 oder in direkter Reihenfolge der einzelnen Seitenzahlen, wobei die aktuelle Seitenzahl markiert ist. Da die Menüs aufgrund der Memorierungskapazität in

markiert ist. Da die Menüs aufgrund der Memorierungskapazität in ihrer Länge relativ begrenzt sind und der Umfang bereits bekannt ist, kann auf die explizite Seitenangabe verzichtet werden. Die Darstellung von Pfeilen ist nicht nur ausreichend, als Symbole sind sie zudem schneller zu interpretieren (ZEIDLER & ZELLNER 1992: 59-60; HERCZEG 1994: 79-80). Sofern das Menü nach oben und / oder nach unten fortgesetzt wird, können entsprechende Richtungspfeile eingeblendet werden. Der Anfang und das Ende des Menüs können durch das Ausblenden der jeweiligen Pfeile implizit ausgedrückt werden.

Verzichtbar ist wiederum die textuelle Darstellung der *Eingabeaufforderung*, da diese im Handlungskontext impliziert ist. Allerdings kann der aktive Zustand des Spracherkenners symbolisiert werden, wobei hier im Gegensatz zum akustischen Modus auch die Dauer der Eingabemöglichkeit angezeigt werden kann. Da i.d.R. die Erkennung nach einer gewissen Zeit automatisch deaktiviert wird, sofern keine Eingabe gemacht wird, erhält der Benutzer die entsprechende Rückmeldung. Auch im Fall der Eingabe ist die Rückmeldung über ein eingegangenes Signal als aufgabenangemessen zu sehen.

Da anzunehmen ist, dass sich eine konsekutive visuelle Menüausgabe unterstützend auf die Bewusstseinshaltung auswirkt, die sprachliche Interaktion als primären Modus zu nutzen, stellt sich weiterhin die Frage, was während der Sprachausgabe auf dem Display erscheint. Um jegliche Ablenkung von der sprachlichen Ausgabe zu vermeiden, wäre ein schwarzes Display zu bevorzugen. Da der aktuelle Systemzustand über den akustischen Modus vermittelt wird, kann der Anspruch der Selbstbeschreibungsfähigkeit als erfüllt gelten. Dennoch könnte eine Verunsicherung hinsichtlich der Displayfunktion auftreten. Ein schwarzes Display lässt nicht erkennen, ob dies intendiert oder einer Störung zuzuschreiben ist. Andererseits könnte durch die Nachrangigkeit der visuellen Ausgabe die Notwendigkeit einer steten Funktionsbestätigung des Displays relativiert werden.

Als Interimsanzeige können drei Möglichkeiten in Betracht gezogen werden: 1) eine Defaultanzeige, 2) ein Symbol, das die Sprachausgabe signalisiert und 3) ein Anfragefeedback. Die *Defaultanzeige* kann, wie ihr Wortlaut bereits impliziert, nur die generelle Funktionalität des Displays bestätigen, wobei sich hier eine Anzeige in symbolischer Form anbietet, wie z.B. das Fahrzeuglogo oder, falls es sich um eine zentrale Schnittstelle mehrerer Systeme handelt, ein Piktogramm, das das aktive System signalisiert. Da der Informationsgehalt eines Symbols mit einem kurzen Blick erfasst werden kann, ist allenfalls von einem minimalen Ablenkungspotential auszugehen. Ein *Symbol*

zur *Signalisierung der Sprachausgabe* hätte neben der schnellen Wahrnehmung zudem den Vorteil, den aktuellen Zustand anzuzeigen. Die textuelle Präsentation des *Anfragefeedbacks* gibt darüber hinaus eine inhaltliche Information, was unter dem Aspekt der Selbstbeschreibungsfähigkeit als optimal angesehen werden kann. Allerdings ist die Verarbeitung textueller Information zeitintensiver als von symbolischer, was von der sprachlichen Ausgabe ablenken könnte. Ferner könnte die Intention der konsekutiven visuellen Menüausgabe, die sprachliche Interaktion als Hauptmodus zu suggerieren, durch die textuelle Darstellung verwässert werden.

4 Versuch

Die diskutierten Optionen wurden in einem realen Fahrversuch untersucht, wobei eine exemplarische Bedienungsanleitung als Anwendungskontext diente.

4.1 Das System

Das System war auf einem Standard-PC unter Windows-NT mit VBA implementiert und befand sich während der Versuche im Kofferraum des Testfahrzeugs. Zu bedienen war es über einen Hebel am Lenkrad. Durch Ziehen konnte die Erkennung aktiviert werden bzw. die Ausgabe unterbrochen werden. Der Hebel diente weiterhin zur Kontrolle des Displays. Nach Beendigung der Sprachausgabe konnte durch ein Drücken nach oben zurückgeblättert, bzw. nach unten weitergeblättert werden. Das Display war oberhalb der Mittelkonsole angebracht. Es handelte sich um ein 7 Zoll TFT-Display mit einer Auflösung von 640 x 480.

Die Informationsabfrage war allein über die Menüsteuerung möglich. Die Menüfolge erstreckte sich maximal über drei Ebenen, der Umfang der Menüs lag zwischen 3 und 14 Menüpunkten. Nach der Systemaufforderung zur Eingabe wurde der Spracherkenner automatisch aktiviert. Die akustische Menüausgabe blieb unabhängig von den visuellen Ergänzungen unverändert.

Als Ergänzungen dienten vier Displayvarianten. Bei zwei Varianten erfolgte die Displaypräsentation des Menüs erst nach Beendigung der Sprachausgabe. Hier wurde während der Sprachausgabe einmal die inhaltslose Defaultanzeige in Form des Fahrzeuglogos dargestellt (*Logo*) und einmal die Anfrage als Feedback visualisiert (*Feedback*). Bei den anderen beiden Varianten wurde das Menü bereits parallel zur Sprachausgabe dargestellt und stand danach

weiterhin zur Verfügung. Die simultanen Menüausgaben erfolgten einmal im Blättermodus (*Paging*) und einmal im Scrollmodus (*Scrolling*). Die visuelle Menüausgabe setzte gleichzeitig zur Ausgabe der akustischen Menüpunkte ein. Bei der Variante *Paging* blieb das Display während der akustischen Ausgabe des Anfragefeedbacks und des Menüumfangs schwarz, während bei der Version *Scrolling* das Anfragefeedback eingeblendet wurde. Ferner wurde bei letzterer der aktuelle Menüpunkt durch Highlighting hervorgehoben, der immer an der zweiten Position des sichtbaren Menüs war. Die visuellen Menüausgaben nach der Sprachausgabe waren für alle Varianten im Blättermodus. Mehrseitige Menüs wurden durch Pfeile an der rechten Seite dargestellt, die in die entsprechende(n) Fortsetzungsrichtung(en) wiesen.

Sprachausgabe	Logo	Feedback	Paging	Scrolling
Anfrage-Feedback	Fahrzeuglogo	Vis. Anfrage-	Black Screen	Vis. Anfrage-
Es gibt x Themen:	(Defaultanzeige)	Feedback	"	Feedback
Menü	"	"	Visuelles Menü;	Visuelles Menü;
Welches Th. wollen	"	"	(Blättermodus)	(Scrollingmodus)
Sie?				
Erkenneröffnung	Visuelles Menü	Visuelles Menü	Visuelles Menü	Visuelles Menü

Tab. 2: Zustände der Systemvarianten

4.2 Der Versuchsablauf

Dem Versuch wurde ein *Within subject design* zugrunde gelegt, durch dessen intrasubjektive Vergleiche Einflussgrößen aufgrund individueller Differenzen reduziert und auf eine Standardisierung verzichtet werden konnte (BORTZ 1993). Zwölf erfahrene Fahrer machten jeweils vier Versuchsfahrten innerhalb von zwei Wochen. Als Strecke diente ein ca. 15 km langer Rundkurs durch die Stadt Regensburg, der durch die Innenstadt, Wohn- und Gewerbegebiete sowie über eine Umgehungsstraße führte und damit die diversen Schwierigkeitsgrade nach FASTENMEIER (1995) erfasste (cf. DAHMEN-ZIMMER ET AL. 1999). In der ersten Fahrt wurde die Versuchsstrecke ohne Systembedienung abgefahren, um eine Kontrollbedingung zu erhalten (F1). Im Anschluss wurden aus Gründen der Kapazität pro Versuchsperson drei der vier Systemvarianten getestet (F2-F4). Um die entstehenden Lerneffekte zu nivellieren, wurde die Reihenfolge der zum Test gegebenen Systemvarianten permutiert, so dass bezüglich der Reihenfolgeposition eine einheitliche Verteilung resultierte.

Pro Fahrt erhielten die Probanden acht Aufgaben, die als klassische Bedienungsanleitungsfragen gelten können. Der Beginn der einzelnen Aufgaben war an bestimmten Streckenabschnitten festgemacht, um die Bearbeitung in

schwierigen Fahrsituationen sicherzustellen (*Testing the limits*). Die Varianten wurden hinsichtlich einer effizienten Systeminteraktion, der subjektiven Belastungssituation und der Verkehrssicherheit überprüft. Dazu dienten Fragebögen für die Probanden, die Beurteilung des Fahrverhaltens durch einen assistierenden Fahrlehrer sowie eine anschließende Analyse der Systeminteraktion anhand von Videoaufzeichnungen.

5 Ergebnisse

Es wurden die Hypothesen aufgestellt, dass 1) die konsekutive Menüausgabe die sprachliche Interaktion als primären Modus besser unterstützt, was zu einer effizienten Systeminteraktion und einem stabilen Fahrverhalten führt; und 2) dass bei primärer sprachlicher Interaktion die Darstellung von redundanten visuellen Metainformationen verzichtbar ist. Da jeder Versuchsperson nur drei der vier Varianten zum Test gegeben wurde, resultierten daraus unterschiedliche Messreihen, wodurch die statistischen Vergleiche über T-Tests für gepaarte Stichproben vorgenommen wurden. Die Bewertungen der Probanden wurden anhand von bipolaren Analogskalen vorgenommen. Auf einer horizontalen Linie markierten sie die für sie zutreffende Position zwischen zwei Polen, z.B. *gut* – *schlecht*. Die Markierungen wurden in Millimetern ausgemessen, was einen Bereich von 0 bis 38 ergab. Die Ergebnisse wurden für die Darstellung auf den Bereich $[0, 1]$ normalisiert.

5.1 Bewertung der Metainformationen

Die **akustischen Metainformationen** wurden von den Probanden für positiv und ausreichend befunden. Die Wiederholung der Anfrage wurde auf der Analogskala im Mittel mit 0,80 bewertet. Die Angabe der Menülänge inform der Ausgabe *Es gibt x Themen* wurde mit durchschnittlich 0,84 bewertet.

Die **visuelle Anzeige des Anfragefeedbacks** erwies sich im Rahmen der Befragung als weniger relevant. Die existente Anzeige wurde zwar relativ hoch bewertet, nämlich mit durchschnittlich 0,68 bei der Variante *Feedback* und mit 0,79 bei der Variante *Scrolling*; die Anzeige wurde jedoch in den anderen Fällen nicht vermisst. Bei den Varianten *Logo* und *Paging* wurde dies im Mittel jeweils mit 0,16 angegeben. Die intrasubjektiven Vergleiche zwischen den Versionen *Feedback* und *Scrolling* ($p = ,136$) sowie zwischen *Logo* und *Paging* ($p = ,287$) waren jeweils nicht signifikant.

Bei den Versionen *Logo* und *Paging* war allerdings eine gewisse Verunsicherung hinsichtlich der Displayfunktionalität zu verzeichnen, sofern der aktuelle **Systemzustand** nicht abzulesen war. Die Verunsicherung durch die Defaultanzeige bei der Version *Logo* betrug durchschnittlich 0,16. Das zeitweilig schwarze Display bei der Version *Paging* wurde diesbezüglich mit 0,24 bewertet. Die mittlere Differenz zwischen den beiden Varianten war nicht signifikant ($p = ,361$).

In einem abschließenden **Ranking von Displayanzeigen**, welche während der Sprachausgabe favorisiert werden würden, wurde ein schwarzes Display dennoch an letzte Position (1,25) gesetzt. An Position 3 war die Defaultanzeige (1,75), an Position 2 eine Statusanzeige, die die aktive Sprachausgabe signalisiert (3,33) und an Position 1 das visuelle Anfragefeedback (3,67).

Das **visuelle Feedback über die geöffnete Erkennung**, das in Form eines roten Balkens am Display erschien, wurde positiv beurteilt. Bei der Version *Logo* betrug die mittlere Bewertung 0,84, bei der Version *Feedback* 0,87, bei der Version *Paging* 0,92 und bei der Version *Scrolling* 0,89. Im intrasubjektiven Vergleich zeigten sich keine signifikanten Differenzen.

	Mittlere Differenz	Stand.- Abw.	Sig. (2-seitig)		Mittlere Differenz	Stand.- Abw.	Sig. (2-seitig)
L – F	-0,08	0,21	,405	F – P	0,02	0,05	,317
L – P	0	0	1,000	F – S	-0,04	0,09	,267
L – S	-0,08	0,31	,534	P – S	0,02	0,08	,595

Tab. 3: Bewertung der Anzeige des aktiven Spracherkenners

Die Anzeige mittels **Pfeilen**, die am rechten Displayrand erschienen, um weitere vorausgehende bzw. nachfolgende Seiten anzuzeigen, wurde von allen außer 1 VP als probat beurteilt. Diese kritisierte, dass die Pfeile nicht auffällig genug waren.

Fazit: Aus Sicht der Probanden kann die Hypothese, dass auf redundante visuelle Informationen verzichtet werden kann, nicht eindeutig bestätigt werden. Die fehlende Anzeige des visuellen Anfragefeedbacks wurde zwar nicht vermisst, doch von allen Optionen, welche Anzeige während der Sprachausgabe erscheinen sollte, favorisiert. Eindeutig zeigte sich dagegen, dass auch im multimodalen Kontext über jede Komponente der Systemzustand abzulesen sein sollte.

5.2 Systeminteraktion

Weiterhin wurde die Systeminteraktion daraufhin untersucht, inwieweit diese von den diversen Varianten beeinflusst wurde. Ein eindeutiges Bild ergab sich hinsichtlich des visuellen Ablenkungspotentials zwischen der konsekutiven und der simultanen Menüausgabe. Die **Blicke während der Sprachausgabe** auf das Display betrugen bei der Version *Logo* durchschnittlich 26,2, bei der Version *Feedback* 39,9, bei der Variante *Paging* 102,7 und bei der Variante *Scrolling* 99,7. Die Differenzen im intrasubjektiven Vergleich zwischen den konsekutiven und den simultanen Varianten waren alle signifikant (cf. Tab. 4). Zwischen den konsekutiven Varianten *Logo* und *Feedback* zeigte sich noch eine marg. signifikante Differenz von 20,2 Blicken. Stellt man die Blickzuwendungen in Relation zu den ausgegebenen Menüpunkten, ergaben sich für die Version *Logo* 0,2 Blicke pro Menüpunkt, bei der Version *Feedback* 0,3 Blicke, bei der Version *Paging* 0,8 Blicke und bei der Version *Scrolling* 0,9 Blicke.

	Blicke während der Spracheingabe			Blicke während der Sprachausgabe			Blicke nach der Sprachausgabe		
	Mittle- re Diff.	Stand.- Abw.	Sig. (2-s.)	Mittle- re Diff.	Stand.- Abw.	Sig. (2-s.)	Mittlere Diff.	Stand.- Abw.	Sig. (2-s.)
L – F	-15,7	33,2	,301	-20,2	20,3	,059	-16,8	29,0	,214
L – P	-7,2	21,9	,459	-77,8	35,5	,003	-2,2	16,6	,762
L – S	-19,7	24,9	,111	-76,2	26,8	,001	6,2	29,3	,628
F – P	-0,5	34,8	,973	-78,8	27,8	,001	10,3	20,6	,275
F – S	4,5	7,3	,192	-63,0	31,0	,004	20,5	16,9	,031
P – S	13,5	36,0	,401	3,0	29,9	,815	-3,7	23,6	,719

Tab. 4: Absolute Displayblicke

Bei der Häufigkeit der Blickzuwendungen auf das Display, **nachdem die Sprachausgabe beendet war**, stellten sich die Differenzen allerdings wieder ein. Pro Fahrt wurden bei der Version *Logo* durchschnittlich 27,9 mal auf das Display gesehen, bei der Version *Feedback* 41,4 mal, bei der Version *Paging* 27,8 mal und bei der Version *Scrolling* 23,1 mal. Die Tendenz der häufigeren Blickzuwendungen bei der Version *Feedback* bestätigte sich im intrasubjektiven Vergleich nur zur Version *Scrolling* (cf. Tab. 4). Im Vergleich der Blickzuwendungen pro erhaltener Menüpunkte traten keine Differenzen auf. Im Mittel beliefen sie sich in der Version *Logo* auf 0,3 Blicke und in den anderen Versionen auf 0,4 Blicke pro Menüpunkt.

Weiterhin war **während der Spracheingabe** bei allen Varianten eine hohe Blickfrequenz zu beobachten. Bei der Version *Logo* waren dies im Mittel 72,2 Blicke, bei der Version *Feedback* 84,0 Blicke und bei den Versionen *Paging* und *Scrolling* jeweils 78,1 Blicke. Setzt man die Blicke in Zusammenhang mit den Benutzereingaben, ergeben sich im Mittel etwa 2 Blicke pro Eingabe. Bei der Version *Logo* waren bei durchschnittlich 36,6 Eingaben 2,0 Blicke zu verzeichnen, bei der Version *Feedback* bei 34,4 Eingaben 2,4 Blicke, bei der Version *Paging* bei 38,7 Eingaben 2,0 Blicke und bei der Version *Scrolling* bei 34,1 Eingaben 2,3 Blicke. Zwischen den Varianten zeigten sich keine signifikanten Differenzen (cf. Tab. 4). Dies deutet darauf hin, dass die bloße visuelle Anzeige über den Erkennerstatus nicht ausreichend war. Im Gegensatz zu einem SIMBA-Versuch, der bei stehendem Fahrzeug durchgeführt wurde (SALMEN ET AL. 1999), war in der Verkehrssituation, in der zahllose visuelle Reize verarbeitet werden, die Aufmerksamkeitsschwelle nicht mehr hoch genug, um aus dem Augenwinkel heraus wahrgenommen zu werden. Folglich ist der Erkennerezustand mit einem zusätzlichen Signal zu indizieren.

Bezüglich der Frage, **wie die Probanden insgesamt mit der sprachlichen Ausgabe zurechtkamen**, ergaben die mittleren Bewertungen bei der Variante *Logo* 0,74, bei der Version *Feedback* 0,85, bei der Version *Paging* 0,85 und bei der Version *Scrolling* 0,80. Die Differenzen zwischen den Varianten erweisen sich alle als nicht signifikant. Auffällig war jedoch, dass in der Kombination von erster Fahrt und Systembedienung nur bei der Variante *Logo* die Bewertungen deutlich niedriger waren, nämlich 0,51. Bei der Version *Feedback*, bei der die Menüausgabe ebenfalls konsekutiv erfolgte, blieben die Bewertungen dagegen konstant. Vergleicht man nur die Bewertungen von Fahrt 3 und Fahrt 4 waren auch die Differenzen der Mittelwerte zwischen den Varianten aufgehoben: *Logo* 0,86, *Feedback* 0,84, *Paging* 0,85 und *Scrolling* 0,82. Folglich trat nur bei der Version *Logo* ein Lerneffekt auf.

	Mittlere Differenz	Stand.-Abw.	Sig. (2-seitig)		Mittlere Differenz	Stand.-Abw.	Sig. (2-seitig)
L – F	-0,08	0,30	,548	F – P	-0,04	0,12	,443
L – P	-0,06	0,15	,352	F – S	0,01	0,11	,776
L – S	-0,04	0,31	,790	P – S	0,07	0,09	,112

Tab. 5: Bewertung des Zurechtkommens mit der Sprachausgabe

Im Vergleich der tatsächlichen **Anfrageschritte** (= exklusive zusätzlicher Anfragen aufgrund systembedingter Fehlerkennungen) mit den minimal benötigten ergab sich für die Version *Logo* eine Differenz von 0,8 ($p = ,006$), bei der Version *Feedback* von 1,0 ($p = ,001$), bei der Version *Paging* von 0,9

($p = ,011$) und bei der Version *Scrolling* von 0,5 ($p = ,007$). Betrachtet man die Differenzen zwischen den tatsächlichen und den minimal benötigten Anfrageschritten nach Einstellen des Lerneffekts, also ohne F2, war diese in der Version *Feedback* weiterhin am höchsten mit 0,8 ($p = ,005$). Bei der Version *Paging* reduzierte sich die Differenz auf 0,6 ($p = ,030$) und bei der Version *Scrolling* auf 0,3 ($p = ,026$). Bei der Version *Logo* war die deutlichste Reduktion mit 0,4 ($p = ,014$) zu beobachten.

Hinsichtlich der **vorzeitigen Menüauswahl**, d.h. vorgenommene Eingaben bevor die Menüausgabe beendet war, zeigte sich erneut, dass mit der Version *Logo* der sprachliche Modus am besten unterstützt wurde. Vorzeitige Eingaben aufgrund der akustischen Wahrnehmung kamen hier durchschnittlich 7,8 mal vor, bei der Version *Feedback* 5,3 mal, bei der Version *Paging* 1,2 mal und bei der Version *Scrolling* 0,7 mal. Bei den simultanen Varianten *Paging* und *Scrolling* wurden vorzeitige Eingaben vornehmlich über das Display gemacht, 8,7 mal bzw. 7,6 mal. Sig. Differenzen bezüglich der akustischen Auswahl erwiesen sich im intrasubjektiven Vergleich nur zwischen den Varianten *Logo* und *Paging* mit 6,83 ($p = ,022$) sowie *Logo* und *Scrolling* mit 8,67 ($p = ,014$).

Die Notwendigkeit der **Menüwiederholungen** erstreckte sich bei der Version *Logo* auf ein durchschnittliches Vorkommen von 2,3, bei der Version *Feedback* von 4,2, bei der Version *Paging* von 3,0 und bei der Version *Scrolling* von 3,1. Im intrasubjektiven Vergleich ergaben sich zwischen den Varianten keine signifikanten Differenzen. Lediglich bei der Version *Feedback* ergab sich eine marg. signifikante Differenz von 2,8 zur Version *Scrolling* ($p = ,077$).

Fazit: Die Hypothese, dass die konsekutive Menüausgabe die sprachliche Interaktion als primären Modus besser unterstützt, kann unter dem Aspekt der Systeminteraktion nur für die Version *Logo* eindeutig bestätigt werden. Da sich bei der Version *Feedback* wie bei den simultanen Varianten kein Lerneffekt bezüglich der sprachlichen Interaktion einstellte, sich keine signifikante Differenz zu den simultanen Varianten bezüglich der vorzeitigen akustischen Menüauswahl einstellte und tendenziell die häufigsten Menüwiederholungen vorkamen, ist davon auszugehen, dass die Aufmerksamkeit auf die sprachliche Ausgabe reduziert war. Dies bestätigt ferner die Annahme, dass die Intention der konsekutiven Menüausgabe, die sprachliche Interaktion als primären Modus zu unterstützen, durch die textuelle inhaltliche Darstellung verwässert wurde. Indirekt wurde dadurch auch die Hypothese, dass auf redundante Informationen verzichtet werden kann, bestätigt.

5.3 Subjektive Belastungssituation

Über eine Befragung der Probanden mittels eines Fragebogens wurde weiterhin untersucht, welche Auswirkungen die Systemvarianten auf die subjektive Belastungssituation haben. Als Vergleichswert diente die Kontrollfahrt.

Der entstandene **Zeitdruck** während der Testfahrt war in der Kontrollbedingung mit einer durchschnittlichen Bewertung von 0,06 sehr niedrig. Bei zusätzlicher Systembedienung war jeweils ein Anstieg zu beobachten, wobei dieser bei der Version *Feedback* nicht signifikant und bei der Version *Logo* lediglich marg. signifikant war (cf. Tab. 6).

	Zeitdruck			Stress		
	Mittlere Differenz	Stand.- Abw.	Sig. (2-seitig)	Mittlere Differenz	Stand.- Abw.	Sig. (2-seitig)
K	0,06			0,14		
L	+0,14	0,19	,060	+0,10	0,15	,089
F	+0,07	0,12	,135	+0,19	0,15	,006
P	+0,18	0,20	,031	+0,24	0,25	,021
S	+0,11	0,11	,015	+0,20	0,22	,025

Tab. 6: Subjektiver Zeitdruck und Stress

Der **Stress** stieg mit zusätzlicher Systeminteraktion ebenfalls an. Während bei der Version *Logo* der niedrigste Anstieg zu verzeichnen war, der zudem nur marg. Signifikanz erreichte, zeigte sich bei der Version *Feedback* und bei den simultanen Varianten ein deutlich höherer Stresszuwachs, der jeweils signifikant war (cf. Tab. 6).

Die **Bewältigung** der Testfahrt nahm bei allen Varianten ab, wobei dies nur bei den Versionen *Feedback* und *Paging* signifikant war. Bei der Version *Logo* war insgesamt der höchste Rückgang zu verzeichnen (cf. Tab. 7). Allerdings resultiert dieser vornehmlich aus den Angaben aus der ersten Fahrt mit Systembedienung. Vergleicht man lediglich F3 und F4 mit der Kontrollfahrt, ist die Differenz mit 0,05 fast aufgehoben ($p = ,467$). Die Werte der anderen Varianten weisen dagegen nur wenig Veränderung auf, wenn sich die Differenzen zur Kontrollfahrt auch nicht mehr als signifikant erweisen: *Feedback*: 0,12 ($p = ,127$), *Paging*: 0,15 ($p = ,137$) und *Scrolling*: 0,13 ($p = ,269$).

Die **Konzentration** ging gegenüber der Kontrollfahrt bei allen Varianten zurück. Da bei der Kontrollfahrt jeweils die besten Werte vorlagen, ist davon auszugehen, dass der Rückgang an Konzentration auf Unkonzentriertheit

verweist. Ein signifikanter Konzentrationsabfall war bei den Varianten *Scrolling* und *Feedback* festzustellen (cf. Tab. 7).

	Bewältigung			Konzentration		
	Mittlere Differenz	Stand.-Abw.	Sig. (2-seitig)	Mittlere Differenz	Stand.-Abw.	Sig. (2-seitig)
K	0,82			0,80		
L	-0,23	0,34	,075	-0,04	0,20	,587
F	-0,17	0,21	,042	-0,12	0,15	,036
P	-0,15	0,17	,034	-0,07	0,16	,255
S	-0,14	0,24	,125	-0,18	0,22	,048

Tabelle 7: Subjektive Bewältigung und Konzentration

Fazit: Da sich bei allen Varianten außer *Feedback* ein erhöhter Zeitdruck einstellte, kann dies als weiteres Indiz dafür gewertet werden, dass hier die Aufmerksamkeit auf die Sprachausgabe reduziert war. Der Stress stieg allerdings wie bei den simultanen Varianten signifikant an, was auf die vermehrte Displayzuwendung zurückzuführen ist, was auch in Unkonzentriertheit resultierte. Die relativ niedrige Bewältigung bei der Version *Logo* in der ersten Fahrt mit Systembedienung geht mit dem Lernprozess bezüglich der sprachlichen Interaktion einher. Damit können auch aus der Perspektive der subjektiven Belastungssituation die Hypothesen bestätigt werden.

5.4 Fahrverhalten

Schließlich wurde der Einfluss der Varianten auf das Fahrverhalten analysiert, wobei die Kontrollbedingung wieder als Vergleichsmaß diente. Ein Kriterium war dabei das verkehrsrelevante Blickverhalten, das in Sicherheits- und Tachoblicke unterteilt wurde. Die **Sicherungsblicke** umfassten Blicke in die Rückspiegel sowie seitliche und rückwärtige Blicke, um die Verkehrssituation bei Spurwechseln oder Abbiegevorgängen zu erfassen. Die höchsten Reduktionen lagen bei den Varianten *Feedback* und *Paging* vor, die jeweils signifikant waren. Der niedrigste Rückgang war bei der Version *Logo* zu verzeichnen, der zudem nicht signifikant war (cf. Tab. 8).

Als **Tachoblicke** wurden Blicke auf die Armaturenanzeigen bezeichnet, da sie meist zur Kontrolle der Geschwindigkeit dienen (cf. DAHMEN-ZIMMER ET AL. 1999). Wiederum zeigten sich die Reduktionen bei den Versionen *Feedback* und *Paging* als signifikant. Der größte Rückgang war dennoch bei der Version *Scrolling* zu verzeichnen, der aber nur marg. Signifikanz erreichte (cf. Tab. 8).

	Sicherungsblieke			Tachoblicke		
	Mittlere Differenz	Stand.- Abw.	Sig. (2-seitig)	Mittlere Differenz	Stand.- Abw.	Sig. (2-seitig)
K	139			112		
L	-15,0	24,9	,108	-36,3	54,8	,082
F	-24,7	31,7	,048	-30,9	18,1	,001
P	-24,4	9,7	,000	-41,1	40,3	,016
S	-20,6	31,5	,086	-49,1	66,6	,058

Tab. 8: Absolute Sicherungs- und Tachoblicke

Das Fahrverhalten erfuh weiterhin durch den assistierenden Fahrlehrer eine Gesamtbeurteilung nach den Kategorien *Sicherheit*, *Kontrolle* und *zügiges Fahren* von 1 (sehr ausgeprägt) bis 5 (wenig ausgeprägt). Bezüglich der Sicherheit stellten sich bei keiner Variante signifikante Differenzen zur Kontrollfahrt ein. Bezüglich der Kontrolle über die Verkehrssituation stellte sich bei der Version *Paging* eine signifikante und bei der Version *Scrolling* eine marg. signifikante Differenz ein. Das zügige Fahren reduzierte sich signifikant bei den Versionen *Logo* und *Paging*.

	Sicherheit			Kontrolle			Zügiges Fahren		
	Mittlere Diff.	Stand.- Abw.	Sig. (2-s.)	Mittlere Diff.	Stand.- Abw.	Sig. (2-s.)	Mittlere Diff.	Stand.- Abw.	Sig. (2-s.)
K	2,3			2,1			2,0		
L	+,22	,83	,447	+,33	,71	,195	+,78	,44	,001
F	+,11	,60	,594	+,22	,83	,447	+,78	1,20	,088
P	+,22	,83	,447	+,44	,53	,035	+,67	,71	,022
S	0	,71	1,000	+,33	,50	,081	+,44	,88	,169

Tab. 9: Gesamtbeurteilung

Ferner wurden während der Fahrt Verstöße im Geschwindigkeitsverhalten festgehalten. Die Verstöße wurden von 1 (leichte Behinderung) bis 6 (Eingreifen des Fahrlehrers) gewichtet und pro Fahrt summiert. Hier zeigte sich bei der Version *Scrolling* ein signifikanter Anstieg, der fast ausschließlich auf Tempoüberschreitungen basierte. Bei der Version *Feedback* zeigte sich ein marg. signifikanter Anstieg, der sowohl auf Tempoüberschreitungen wie -unterschreitungen basierte, in der Unterscheidung jedoch keine Signifikanz mehr erreichte (cf. Tab. 10).

Fazit: Insgesamt zeigte sich bei keiner Variante eine signifikante Reduktion der Sicherheit. Allerdings ließen sich qualitative Unterschiede beobachten.

Bei der Version *Logo* war lediglich ein Rückgang des zügigen Fahrens zu verzeichnen. Die Version *Feedback* nimmt dagegen wiederum eine Zwischenstellung ein. Die Sicherungsblicke wurden reduziert, aber im Gegensatz zur Version *Paging* blieb die Kontrolle über die Verkehrssituation aufrechterhalten. Die reduzierten Tachoblicke führten nur zu einer marg. signifikanten Differenz im Geschwindigkeitsverhalten, während bei der Version *Scrolling* trotz marg. signifikant reduzierter Tachoblicke mehr Tempoüberschreitungen vorkamen. Insofern kann die Hypothese, dass eine primäre sprachliche Systeminteraktion ein stabiles Fahrverhalten ermöglicht wiederum nur für die Version *Logo* bestätigt werden.

	Geschwindigk.-verstöße			G-Tempoüberschreit.			G-Tempounterschreit.		
	Mittle- re Diff.	Stand.- Abw.	Sig. (2-s.)	Mittle- re Diff.	Stand.- Abw.	Sig. (2-s.)	Mittlere Diff.	Stand.- Abw.	Sig. (2-s.)
K	10,1			9,5			0,6		
L	+2,33	4,18	,133	+1,00	2,69	,298	+1,22	2,64	,202
F	+2,67	3,84	,071	+,78	3,46	,519	+1,78	4,02	,222
P	+1,44	5,15	,425	-1,11	1,96	,128	+2,67	4,74	,130
S	+2,56	2,74	,023	+2,11	2,52	,036	+,44	1,01	,225

Tab. 10: Geschwindigkeitsverhalten

6 Fazit

Der Anspruch an das Systemdesign, dem Benutzer die primäre Nutzung der Sprache als Interaktionsmodus zu suggerieren, kann nur durch eine konsekutive visuelle Menüausgabe erreicht werden. Um die volle Konzentration auf die Sprachausgabe zu lenken, muss auf jegliche textuelle Darstellungen auf dem Display verzichtet werden. Bereits die Anzeige des Anfragefeedbacks bedeutet eine Ablenkung, die in einer ineffizienteren Systeminteraktion resultiert und sich in der Konsequenz negativ auf das Fahrverhalten auswirkt. Entsprechend sind alle Metainformationen, die über den akustischen Kanal vermittelt werden können, sprachlich oder akustisch auszugeben, um kontinuierlich in diesem Modus agieren zu können. Redundante visuelle Metainformationen sind für eine effiziente Systeminteraktion nicht erforderlich. Als relevant erwies es sich allerdings, dass über alle Komponenten der aktuelle Systemzustand jederzeit wahrnehmbar ist, um Verunsicherungen der Benutzer zu vermeiden. Dabei sind symbolische Anzeigen zu präferieren, da sie schnell erfassbar sind und dadurch die akustische Informationsverarbeitung nicht beeinträchtigen.

7 Literatur

- Balentine, B. 1999. "Re-engineering the Speech Menu. A 'Device' Approach to Interactive List-Selection." In: D. Gardner-Bonneau (ed.). 1999. Human Factors And Voice Interactive Systems. Kluwer Academic Publishers. 205-35.
- Ballas, J.A. und T. Mullins. 1991. "Effects of Context on the Identification of Everyday Sounds." Human Performance 4 (3). 199-219.
- Becker, S., M. Brockmann, E. Bruckmayr, O. Hofmann, R. Krause, A. Mertens, R. Nin und J. Sonntag. 1995. Telefonieren am Steuer. Bergisch Gladbach: Wirtschaftsverlag NW.
- Bengler, K. 1995. Gestaltung und experimentelle Untersuchung unterschiedlicher Präsentationsformen von Wegleitungsinformationen in Kraftfahrzeugen. Regensburg: S. Roderer Verlag.
- Bortz, J.. 1993. Statistik für Sozialwissenschaftler. Berlin u. Heidelberg: Springer Verlag.
- Brewster, S.A., P.C. Wright und A.D.N. Edwards. 1994. "The Design and Evaluation of an Auditory-Enhanced Scrollbar." In: Human Factors in Computing Systems. CHI '94 Conference Proceedings. New York: ACM Press. 173-79.
- Dahmen-Zimmer, K., M. Kostka, W. Piechulla, I. Scheufler und A. Zimmer. 1999. "KOMI-ZIF. Kompendium für dedizierte Methoden bei der Untersuchung von informationellen Zusatztätigkeiten im Fahrzeug." Projektbericht. Universität Regensburg.
- De Vries, G. und G.I. Johnson. 1997. "Spoken Help for a Car Stereo: An Explanatory Study." Behavior & Information Technology 16 (2). 79-87.
- Färber, B. 1987. Geteilte Aufmerksamkeit. Grundlagen und Anwendung im motorisierten Straßenverkehr. Köln: Verlag TÜV Rheinland.
- Fastenmeier, W. 1995. "Die Verkehrssituation als Analyseeinheit im Verkehrssystem." In: W. Fastenmeier (ed.). 1995. Autofahrer und Verkehrssituation. Neue Wege zur Bewertung von Sicherheit und Zuverlässigkeit moderner Strassensysteme. Köln: Verlag TÜV Rheinland. 27-78.
- Franz, E.A., H.N. Zelaznik und A. Smith. 1992. "Evidence of Common Timing in the Control of Manual, Orofacial, and Speech Movements." Journal of Motor Behavior, 24. 281-87.
- Hasebrook, J. 1995. Multimedia Psychologie. Eine neue Perspektive menschlicher Kommunikation. Heidelberg u.a.: Spektrum Akademischer Verlag.
- Herczeg, M. 1994. Software-Ergonomie. Grundlagen der Mensch-Computer-Kommunikation. Bonn u.a.: Addison-Wesley.
- Kozma, R.B. 1991. "Learning with Media." Review of Educational Research 61 (2). 179-211.
- Ludwig, M. 1998. "Auswertung der Benutzertests des 1. Gesamtprototyps der interaktiven Betriebsanleitung." Projektbericht, Universität Regensburg.
- Ludwig, M., P. Großmann und A. Salmen. 1996. "Test des Erstbetriebmodus." Projektbericht, Universität Regensburg.
- Miller, G.A. 1956. "The Magical Number Seven, Plus Or Minus Two: Some Limits on Our Capacity for Processing Information." Psychological Review 63 (2). 81-97.

- Mullet, K. und D. Sano. 1995. Designing Visual Interfaces. Communication Oriented Techniques. Mountain View, CA.:Sun Soft Press.
- Pashler, H. und J.C. Johnston. 1998. "Attentional Limitations in Dual Task Performance." In: H. Pashler (ed.). 1998. Attention. East Sussex: Psychology Press. 155-89.
- Salmen, A. und C. Tauschek. 1999. "Auswertung der Elba-Pretests (Prototyp II)." Projektbericht, Universität Regensburg.
- Salmen, A., C. Tauschek, T. Schnakenberg und B. Wisbauer. 1999. "Auswertung der Elba-Benutzertests (Prototyp III)." Projektbericht, Universität Regensburg.
- Simon, H.A. 1974. "How Big Is a Chunk?" Science 183. 482-88.
- Zeidler, A. und R. Zellner. 1992. Software-Ergonomie. Techniken der Dialoggestaltung. München u. Wien: R. Oldenbourg Verlag.



Effects of Manual Phonetic Transcriptions on Recognition Accuracy of Streetnames

Thomas Bauer¹, Ludwig Hitzenberger¹, Marcus Hennecke²

¹Information Science
University of Regensburg
D-93040 Regensburg
Germany

{Thomas.Bauer, Ludwig.Hitzenberger}
@sprachlit.uni-regensburg.de

²Temic Speech Processing
Söflinger Str. 100
D-89077 Ulm
Germany

marcus.hennecke@temic.com

Abstract

In automotive systems, the accurate recognition of street- and place names as needed e.g. for speech controlled navigation systems is difficult, since many of them do not follow regular transcription rules. We report on experiments investigating the effects of correct manual transcription versus strictly automatic transcription, with regard to recognition accuracy. Our experiments showed that in case of large recognizer lexica, the recognition rate could be increased by about 10% by using a manually corrected versions of the phonetic transcriptions used for the recognition process.

1 Introduction

In automotive systems, speech recognition and dialogue systems are becoming more and more important. At the same time navigation systems belong to the standard equipment in cars and, obviously used while driving, speech seems to be the only adequate means of interaction[1]. These systems need to handle an enormous amount of data. There are about 60000 place names and more than half a million street names and points of interest in Germany. Proper names are typically a problem in speech -recognition [4], especially if automatic transcription is used since they often do not follow regular transcription -rules. Streetnames are even worse, with many of them combining words from different languages. In Germany we often see names like John-F.-Kennedy-Platz or Clermont-Ferrand-Allee. Another problem is the lack of



context for the disambiguation which we normally encounter in other dialogue-system domains. There is also a very flat hierarchy of the search space (e.g. the lowest level of the street name list is the concerned city). Thus, the recognition rate on the word level is crucial for the success of the dialogue. One way to improve the recognition rate is supposed to be the adequate phonetic transcription of names with respect to the monolingual recognition system. The transcription of the vocabulary items typically follows static rules applied through an automatic transcription device. On the other hand manual transcription is very costly and time consuming, but so far the only way to achieve more correct transcriptions and thus better recognition rates. In this paper we report on experiments designed to measure the achievable improvements using manual transcriptions over automatic transcriptions.

2 Experimental setup

2.1 Transcriptions

For our experiments we used a data set consisting of the street names of five big German cities (Munich, Stuttgart, Sindelfingen, Ingolstadt, and Regensburg) with a total of 11437 street names in the official orthographic representation taken from a standard navigation database provided by Navigation Technologies (NavTech)¹. These items were transcribed automatically using the internal automatic transcription of the lexicon tool of the Temic speech-recognizer [2]. The output was subsequently checked by a phonetic expert and corrected if necessary. This correction was supported by an automatic transcription tool that uses a learning algorithm implemented by Sympalog². About 15% of the data had to be corrected manually. The system, trained by these corrections, subsequently changed another 50% of the automatic transcriptions. The transcription alphabet used was a slightly modified SAMPA representation.

¹ NavTech is a company providing global geographical navigation data. Further information can be found at www.navtech.com.

² Sympalog is a company providing dialogue-systems and tools. Further information can be found at www.sympalog.com.

2.2 Vocabularies

For the tests we produced four different sets of vocabulary using the lexicon tool of the Temic recognizer. Thus we have two independent variables, the vocabulary size and the correctness of the phonetic transcriptions.

Name	Number of items	Description
Vok1	11437	Automatic transcription without manual corrections.
Vok2	11437	Automatic transcription including manual corrections.
Vok1_s	110	Automatic transcription without manual corrections.
Vok2_s	110	Automatic transcription including manual corrections.

Table 1: Vocabularies

2.3 Test data

A random set of 110 street names was spoken and recorded by 34 test persons. We had 17 female and 17 male speakers with an average age of 29 years. These test persons used a recording tool and recorded the names unsupervised. The audio format was 44100Hz/16Bit. Afterwards the recorded samples were converted to 8000Hz/16Bit PCM raw data (the proper format for the recognizer) using CoolEdit pro 1.1.

1.4 Speech recognition

The test data were processed by the Temic speech recognizer³ [3] on a WindowsNT 4.0 platform using the different vocabularies (Vok1, Vok2, Vok1_s, Vok2_s). The recognizer was activated by Microsoft Access using VBA. The recognition results were stored in this Access database.

2 Results

2.1 Recognition accuracy

	Vok1	Vok2	Vok1_s	Vok2_s
No. of samples tested	3740	3740	3740	3740
No. of samples recognized correctly	2763	3127	3374	3600
Recognition Rate [%]	73.9	83.6	90.2	96.3

³ Temic Star Rec DSR 1.7.

Table 2: Recognition rates

The table 2 shows the recognition rates of the four different vocabularies. The number of samples is the total number of all utterances in the test data. Since each test person recorded all the street names, this sums up to 3740 samples. The following box plot shows the distribution of the recognized items and the appropriate confidence interval (95%). The grey box represents the second and third quartile of the data, while the line in the box indicates the median of the data.

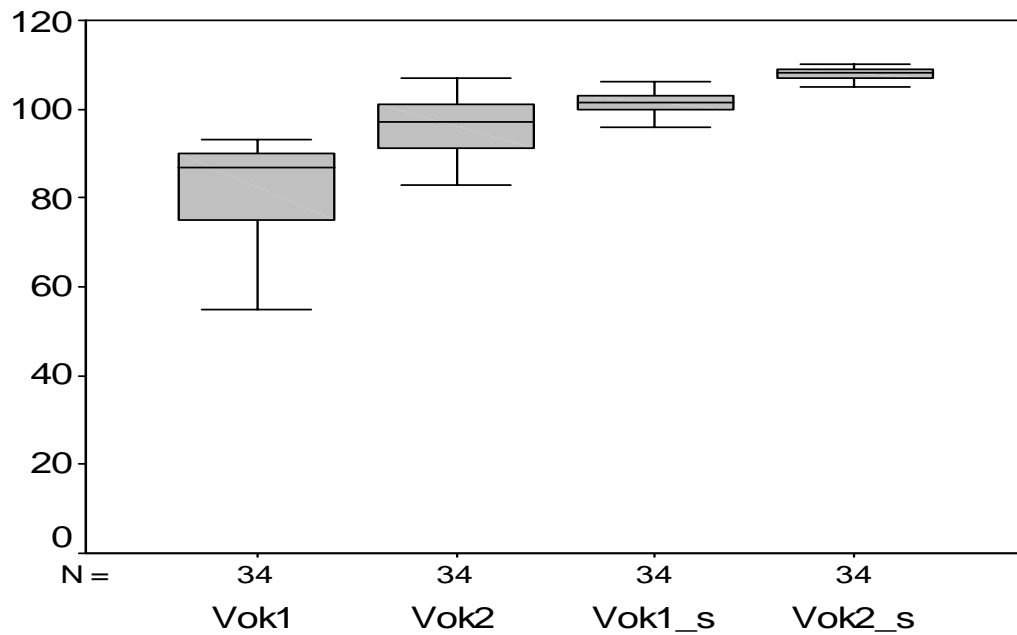


Diagram 1: Distribution of absolute number of correctly recognized items and corresponding vocabulary

2.2 Significance

As table 3 shows, all of the tests were highly significant.

Pairs	Mean of the difference	Standard deviation	Standard error of the mean	95% confidence interval of the difference		Significance
				lower	upper	
Vok1 Vok2	-10.71	3.52	.60	-11.93	-9.48	.000
Vok1 Vok1_s	-17.97	9.19	1.58	-21.18	-14.77	.000
Vok1 Vok2_s	-24.62	9.58	1.64	-27.96	-21.28	.000
Vok2 Vok1_s	-7.26	10.13	1.74	-10.80	-3.73	.000
Vok2 Vok2_s	-13.91	10.01	1.72	-17.40	-10.42	.000
Vok1_s Vok2_s	-6.65	2.56	.44	-7.54	-5.75	.000

Table 3: T-Test of paired samples

3 Discussion

The recognition results show a very strong effect of correct transcriptions in the vocabulary of the recognizer. For the larger vocabulary we achieved an improvement from 73.9% to 83.6. This is a difference of 9.7%. For the small vocabulary there was an improvement from 90.2 to 96.3%, which is a difference of 6.1%. This indicates that the effect is stronger on larger vocabularies; which will be subject of further investigations.

4 Conclusions

Although it is very time consuming and costly to manually correct the transcriptions of large vocabularies, in the context of navigation data and proper names like street and place names there is an obvious benefit in doing so. This is exactly what our experiments show. Also the learning transcription algorithms benefit from the correction of the data, so that the intellectual effort will gradually diminish, but a certain amount of manual control will always be necessary.

5 References

- Hitzenberger, Ludwig: Man Machine Interaction in Car Information Systems, LREC'98, Vol.1:179-182, 1998.
- Lehtinen, G. et. al. IDAS: Interactive Directory Assistance Services, Voice Operated Telecom Services 2000, Ghent.
- StarRecTM GDS (Grammar Development System). TEMIC Telefunken microelectronic GmbH, Ulm, Germany, 2000.
- StarRecTM DSR (DaimlerChrysler Speech Recogniser). TEMIC Telefunken microelectronic GmbH, Ulm, Germany, 2000.
- Lehtinen, G. et. al. IDAS: Interactive Directory Assistance Services, Voice Operated Telecom Services 2000, Ghent.



Entwicklung eines Modells zur Akzeptanzproblematik von Informationstechnologie

Elisabeth Milchrahm

Karl-Franzens-Universität Graz
Institut für Informationswissenschaft
Universitätsstraße 15
A-8010 Graz
elisabeth.milchrahm@kfunigraz.ac.at

Zusammenfassung

Die Akzeptanzforschung von Informationstechnologie hat in den letzten Jahren aufgrund der rasanten Entwicklung des Internets einen neuen Aufschwung erfahren. Konstituierende Elemente der Akzeptanzforschung sind benutzerorientierte Eigenschaften sowie aufgabenspezifische und systemcharakteristische Merkmale. Gegenstand der vorliegenden Arbeit ist die Analyse der Einflussfaktoren auf die Nutzerakzeptanz von Informationssystemen. Anhand multivariater Analysemethoden wird ein Akzeptanzmodell entwickelt und im Rahmen einer repräsentativen Studie zur Akzeptanz von Informationssystemen im bankbetrieblichen Risikomanagement getestet. Im Endergebnis steht ein statistisch signifikantes Akzeptanzmodell, das zu unmittelbaren Konsequenzen in der Praxis geführt hat.

1 Problemstellung

Eine Analyse wirtschaftswissenschaftlicher Publikationen zeigt, dass sich „Akzeptanz“ als Synonym für ein bestimmtes (Nutzungs-)Verhalten weitgehend durchgesetzt hat [VeDa00; GoRe00]. Der Erfolg von Informationstechnologie entscheidet sich vorwiegend durch Art und Ausmaß der Akzeptanz. Dies bedeutet einen Wechsel von einer dichotomen Ausprägung der Akzeptanz im Sinne einer Ja/Nein-Grundsatzentscheidung, wie es von der klassischen Akzeptanztheorie postuliert wird, hin zu einer dynamischen Sichtweise: Dabei wird der Versuch unternommen, Akzeptanz als Nutzungskontinuum zu betrachten, wobei eine tendenziell hohe Nutzungsintensität als hohe Akzeptanz bzw. eine tendenziell geringe Nutzungsintensität als niedrige Akzeptanz



interpretiert wird [vgl. Koll00, 71]. Demzufolge bezieht sich der Akzeptanzbegriff im Rahmen der vorliegenden Arbeit auf die wiederkehrende Entscheidung des Anwenders zur häufigen aufgabenspezifischen Nutzung des Informationssystems. Akzeptanzprobleme äußern sich in der Regel in der unterbleibenden bzw. nicht adäquaten Nutzung der Informationstechnologie. Für den Nutzer bedeuten Akzeptanzprobleme, dass seine Vorstellungen über die Aufgabenerfüllung unter Verwendung der Informationstechnologie nur mangelhaft Berücksichtigung finden [vgl. Hilb84, 321].

Wenngleich in der Literatur Einigkeit über die Symptome der Akzeptanz besteht, so bleibt die Frage nach den Ursachen von Akzeptanz bzw. Akzeptanzproblemen weitgehend offen. Für eine anwendungsorientierte Entwicklung von Informationstechnologie sind jedoch Kenntnisse über die akzeptanzbeeinflussenden Bestimmungsgrößen unabdingbar. Die empirische Akzeptanzforschung ist daher in engem Zusammenhang mit der Implementierungsforschung zu sehen [vgl. Meye99, 116]. Dieser Argumentation Folge leistend, ist es die Intention der vorliegenden Arbeit, die wesentlichen Einflussgrößen auf die Akzeptanz von Informationssystemen (IS) zu untersuchen. Bisherige Ansätze zur Akzeptanzforschung im Bereich der Informationssysteme beschränken sich vorwiegend auf zwei Einflussfaktoren: Die Benutzerfreundlichkeit des Systems sowie die Qualität der zur Verfügung gestellten Informationen sind die Hauptdeterminanten dieser Modelle. Diese Bestimmungsgrößen liefern jedoch einen zu geringen Erklärungsanteil zur Akzeptanz von Informationssystemen. Als Erweiterung der vorgestellten Ansätze wird daher im Rahmen der vorliegenden Arbeit ein neu entwickeltes Akzeptanzmodell empirisch überprüft. Die Studie untersucht die Akzeptanz von Informationssystemen zur Unterstützung von gut strukturierten betrieblichen Problemstellungen ohne hohe Änderungshäufigkeit. Dieser Aufgabenbereich ist neben weiteren Aspekten, wie die regelmäßige Generierung von Standardberichten, ein wichtiges Merkmal von Managementinformationssystemen.

2 Zielsetzung und Konzeption der Untersuchung

Hauptintention der vorliegenden Arbeit ist, die wesentlichen Einflussfaktoren auf die Akzeptanz von Informationssystemen anhand eines entwickelten hypothetischen Kausalmodells zu untersuchen. Dabei soll die neu eingeführte Modellkomponente SYSTEMVERTRAUEN auf ihre Bedeutung zur Klärung des Akzeptanzmodells statistisch getestet werden.

Die empirische Überprüfung des Akzeptanzmodells basiert auf einer Grundgesamtheit von 1200 Mitarbeitern der Bayerischen HypoVereinsbank AG

(HVB) im deutschsprachigen Raum, die als Zielgruppe des konzernweit neu implementierten Informationssystems verstärkt in den Informationsverarbeitungsprozess der Kreditbearbeitung eingebunden sind. Alle Mitarbeiter der Zielgruppe aus den Bereichen Privat- und Geschäftskunden (PuG) sowie Firmenkunden (FKD) verfügen über einen Arbeitsplatz-Computer mit Internet-Zugang.

Die methodische Vorgehensweise der Untersuchung gliedert sich in eine Vor-, Haupt- und Nachstudie, wobei die Vorstudie bzw. Nachstudie anhand von 58 persönlichen Interviews durchgeführt wurden. Als Erhebungsmethode für die Hauptstudie (20.6 - 31.8.2000) wurde ein Intranet-Fragebogen - gekoppelt an eine Excel-Datenbank - gewählt. Der ordinalskalierte Fragebogen wurde dabei auf der Startseite des konzernweiten Informationssystems der HVB als Hyperlink positioniert. Ein entsprechender Begleittext zielte darauf ab, die Motivation der Mitarbeiter zur Mitwirkung an der Umfrage zu steigern. Die Untersuchung stützt sich auf eine Rücklaufquote von 45 % (542 Antworten). Die Datenauswertung erfolgte mit dem Softwareprogramm SPSS (SPSS Inc., Chicago, Il.).

3 Ansätze zur Akzeptanzforschung

Wissenschaftler aus dem Bereich Managementinformationssysteme arbeiten zur Akzeptanzforschung auf Basis von Theorien aus der Sozialpsychologie. Davis überträgt diese theoretischen Grundlagen auf die konkrete Anwendungssituation einer IS-Nutzung [Davi86]. Im Rahmen des „Technology Acceptance Model“ (TAM) geht der Autor von der Annahme aus, dass die Akzeptanz eines IS von der Absicht des Nutzers bestimmt wird, das IS auch in Zukunft zu verwenden („Intention to Use“). Diese Nutzungsabsicht wird geprägt von der Einstellung des Nutzers zum Informationssystem („Attitude toward Using“) und seiner Einschätzung betreffend der Nützlichkeit des IS („Usefulness“). Neben dem Konstrukt „Usefulness“ ist die Benutzerfreundlichkeit bzw. der Faktor „Ease of Use“ eine weitere zentrale Bestimmungsgröße des Akzeptanzmodells. Externe Variable wie beispielsweise die organisatorische Umsetzung von Informationstechnologie, nehmen indirekt Einfluss auf die Akzeptanz des Systemnutzers.

Davis et al. unterzogen dieses theoretische Modell in Folge einer ersten empirischen Überprüfung und kamen zu dem Ergebnis, dass die statistischen Daten das ursprüngliche TAM nur teilweise unterstützen [DaBa89]. Unterschiede zum theoretischen Modell ergeben sich vor allem aus dem Fehlen der Konstrukte „Attitude toward using“ und „External Variables“, die beide keinen

signifikanten Beitrag zur Erklärung des Akzeptanzmodells leisten (vgl. auch [RoHe00, 441; VeDa00, 195]). Die geringe Bedeutung des Konstruktes „External Variables“ im Rahmen des TAM ist insofern von Interesse, als das andere Studien den Einfluss externer Faktoren für einen erfolgreichen IS-Einsatz bestätigen [MaSn89]. Diese Studien gehen jedoch in ihren Untersuchungen von anderen Prämissen aus. Die Ergebnisse der von Davis et al. durchgeführten Untersuchung bestätigen, dass das Konstrukt „Intention to Use“ die zentrale Größe zur Bestimmung des Nutzerverhaltens ist.

4 Modellkomponenten

Die im vorliegenden Akzeptanzmodell untersuchten Faktoren werden als wesentliche Erfolgsgrößen zur Bestimmung der Akzeptanz von IS angesehen. Im Folgenden werden die Konstrukte des Modells in Beziehung zur Akzeptanz des Systemnutzers gesetzt und ihre Bedeutung im Rahmen der Nutzung von IS aufgezeigt. Aufgrund ihres hohen Erklärungsanteils zur Akzeptanzfragestellung werden vor allem die zwei Faktoren SYSTEMVERTRAUEN¹ und EASE OF USE näher erläutert.²

4.1 Konstrukt Systemvertrauen

Zur Charakterisierung der im Rahmen der vorliegenden Studie neu entwickelten Modellkomponente SYSTEMVERTRAUEN gab es keine Möglichkeit auf bereits empirisch getestete Indikatoren zurückzugreifen. Die Vertrauensproblematik im Umfeld von IS wird in der Literatur erst seit kurzer Zeit und nur in Ansätzen thematisiert. Im Mittelpunkt der Untersuchungen steht dabei das Vertrauen in die technische Sicherheit elektronischer Märkte [Ratn98; Reag96, Kief00]. Im Rahmen dieser Schwerpunktsetzung finden informationswissenschaftliche Fragestellungen keine ausreichende Berücksichtigung.³

Eine wichtige Variable die zur Operationalisierung des Konstruktes SYSTEMVERTRAUEN herangezogen wird, ist der Indikator “Nachvollziehbarkeit der Vorgaben”: Das Vertrauen des Nutzers in ein IS wird positiv beeinflusst,

¹ Im Gegensatz zu den anderen Bezeichnungen der Konstrukte wird bewußt auf eine englische Übersetzung verzichtet, um die Weiterentwicklung des Akzeptanzmodells um den vorliegenden Faktor hervorzuheben.

² Weitere Ausführungen zu den Konstrukten würden den Rahmen dieser Arbeit sprengen.

³ In der neueren wirtschaftswissenschaftlichen Literatur werden unter dem Begriff „e-trust“ Möglichkeiten zur Realisierung vertrauensbildender Maßnahmen in elektronischen Märkten andiskutiert.

wenn die veröffentlichten Handlungsanweisungen vom Nutzer nachvollziehbar sind. Als Handlungsanweisungen werden verbindliche Vorgaben und Bearbeitungshinweise verstanden, die dem Nutzer die Unternehmenspolitik vermitteln. In diesem Zusammenhang betont bereits Ackoff, dass komplexe Entscheidungsprozesse entsprechende Entscheidungsregeln erfordern [vgl. Acko67, 150]. Zur Unterstützung der Entscheidungsfindung ist eine bloße Informationsbereitstellung ohne Handlungsempfehlungen nicht ausreichend.

Dem Indikator "Relevanz der Vorgaben" liegt die Annahme zugrunde, dass SYSTEMVERTRAUEN gefördert wird, wenn der Nutzer von der Relevanz der veröffentlichten Handlungsanweisungen überzeugt ist. Im Sinne eines pragmatischen Informationsverständnisses [Wers74; Kuhl89]⁴ wird vor allem die Anforderung an Information gestellt, handlungsrelevant, d.h. einschlägig für die aktuelle Situation sein zu müssen. Kühlen diskutiert das Kriterium der Handlungsrelevanz im Kontext von informationeller Unsicherheit: in welchem Ausmaß wird den zur Verfügung gestellten Informationen getraut, und wie werden die Informationen in ihrer Handlungsrelevanz beurteilt. Dieses sogenannte Validitätsproblem ist ein Problem des Vertrauens [vgl. Kuhl99, 179 ff.].

4.2 Konstrukt Ease of Use

Diese Modellkomponente beruht auf der Annahme, dass ein Zusammenhang zwischen einer nutzerorientierten IS-Gestaltung und der Akzeptanz eines IS existiert.

Zur Gestaltung von IS finden sich in der Literatur zahlreiche Empfehlungen [Alte99]. Ziel ist es, das IS benutzerfreundlich und aufgabenangemessen zu gestalten. Dies ist dann der Fall, wenn das IS die Erledigung einer Arbeitsaufgabe des Nutzers unterstützt, ohne ihn durch Eigenschaften des IS unnötig zu belasten. Ein gut strukturierter IS-Aufbau und entsprechende Suchmöglichkeiten sind dafür die Grundvoraussetzung. Darüber hinaus ist es für die Akzeptanz des IS entscheidend, dass dem Nutzer Informationen nicht nur in angemessener Art, sondern auch in angemessener Qualität und Menge zur Verfügung gestellt werden [Alex98, 38f.]. Ein wichtiges Akzeptanzkriterium von IS ist in diesem Zusammenhang die Bildung informationeller Mehrwerte. In Anlehnung an die Wertschöpfungskette von Porter [Porte85] werden dabei die Prozesse von der ursprünglichen

⁴ Zum Informationsbegriff gibt es zahlreiche Publikationen. In der klassischen Informationstheorie dominiert die technische Sichtweise [ShaWe49], während die in der vorliegenden Arbeit beschriebene informationswissenschaftliche Sichtweise die pragmatische Relevanz von Information betont.

die Prozesse von der ursprünglichen Produktion einer Wissensseinheit bis hin zur Informationsleistung auf ihre Wertschöpfungsanteile untersucht [vgl. Kuhl95, 81]. Von besonderer Bedeutung für die Akzeptanz von IS ist die Nutzersicht von informationellen Mehrwerten: IS generieren dann informationelle Mehrwerte, wenn der einzelne Nutzer zum Beispiel Verbesserungen hinsichtlich seiner Arbeitseffizienz erzielen kann. Diese Mehrwerte mit Effizienzwirkung ermöglichen, dass bisherige Tätigkeiten - wie beispielsweise die Informationsbeschaffung zur Kreditbearbeitung - durch die Nutzung eines IS schneller und leichter durchgeführt werden können.

Auf die Situation der Informationsvermittlung mittels eines IS trifft folgende, von Schmidt als „synoptische, verdichtende Informationsaufbereitung mit (...) integrativen Mehrwerteffekten“ beschriebene Form der Informationsaufbereitung zu [Schm92, 40 f.]: Ziel dieser Informationsvermittlung ist es primär, die Informationen in einem größeren Zusammenhang darzustellen. Vergleichende Analysen von ausgewählten Unternehmen in Kernbranchen der Bank sind hier als Beispiel zu nennen. Eine Möglichkeit der synoptischen Informationsaufbereitung ist die Verdichtung von Informationen, beispielsweise durch Bereitstellung von Überblickskapitel, die den Nutzer in ausreichender Weise über den Inhalt der Originalkapitel informieren. Zur Erarbeitung qualifizierter Entscheidungen auf Basis eines IS ist es erforderlich, dass genaue Kenntnisse des Informationsverarbeitungsprozesses vorliegen. Mit einem unternehmens-weiten IS, beispielsweise im Kreditrisikobereich einer Bank, werden dem Nutzer Informationen bereitgestellt, die ihn bei der Abwicklung von Kreditgeschäften - vom Kreditantrag bis zur Kreditgewährung - begleiten. Der gesamte Kundenmanagementprozess soll durch das IS abgebildet werden. Dabei ist es wichtig, den Systemnutzer durch möglichst vollständige Informationen bei der Informationsbeschaffung im Rahmen der Kundenbetreuung zu entlasten. Diese inhaltliche Betreuung von IS setzt entsprechende Fachkenntnisse der verantwortlichen Informationsvermittler voraus.

4.3 Konstrukt Usefulness

Der Faktor USEFULNESS ist eine wichtige Bestimmungsgröße im Rahmen des TA-Modells. Empirische Studien belegen den Einfluss des Konstruktes USEFULNESS auf die Akzeptanz von IS [LeMa00; LiLu00]. Die vorliegende Modellkomponente geht von der Annahme aus, dass es einen positiven Zusammenhang zwischen der Qualität der zur Verfügung gestellten Informationen und der Akzeptanz von IS gibt [vgl. Gati94, 127]. Dieser Argumentation folgend, wird unter der Modellkomponente USEFULNESS die Bedeutung des IS zur allgemeinen Entscheidungsunterstützung im Unternehmen verstanden.

DeLone/MacLean identifizieren auf Basis einer Meta-Analyse von 180 wissenschaftlichen Arbeiten die wichtigsten Bestimmungsgrößen für einen erfolgreichen IS-Einsatz [DeMc92]. Dabei kommen die Autoren zu dem Ergebnis, dass die Qualität von Information eine wichtige Erfolgsgröße ist. Zur Beurteilung der vom System zur Verfügung gestellten Informationen wird das Kriterium der Relevanz als wichtiger Indikator von allen untersuchten empirischen Arbeiten bestätigt.

Bei der Wahl von Indikatoren zur Operationalisierung von Informationsqualität ist vor allem zu berücksichtigen, welche Zielsetzung das zu untersuchende IS verfolgt. Wie bereits eingangs erwähnt wurde, liegt der Aufgabenschwerpunkt von Managementinformationssystemen in der Vermittlung von relevanten Informationen zur Unterstützung strukturierter Aufgabengebiete ohne hohe Änderungshäufigkeit. Diese Schwerpunktsetzung auf die Dimension der Relevanz von Information entspricht den Ergebnissen empirischer Untersuchungen im Forschungsbereich TAM [MoKi01; LeMa00]. Die Relevanz von Information wird daher als zentrale Messgröße der Modellkomponente USEFULNESS herangezogen.

4.4 Konstrukt Intention to Use

Im Rahmen des Forschungsbereiches TAM wird die Akzeptanz von IS auf Basis der wiederkehrenden Entscheidung zur häufigen Systemnutzung gemessen. Unter dem Konstrukt INTENTION TO USE wird daher die Absicht des Nutzers verstanden, das IS auch in Zukunft zu verwenden. Im gegenwärtigen Forschungsbereich TAM wird vorwiegend die Post-Implementierungsphase von IS als Untersuchungszeitraum zur Erhebung der Akzeptanz herangezogen. Darüber hinaus etabliert sich in der Akzeptanzforschung immer stärker die Vorgangsweise, das Konstrukt „Intention to Use“ näher zu untersuchen, die Beziehung zu dem Faktor „Actual System Use“ jedoch bei der weiteren Analyse auszuklammern. [vgl. Math91, 186] Diese Vorgangsweise ist mehrfach und in hohem Maße durch Resultate empirischer Studien abgesichert: Venkatesh/Davis [vgl. VeDa00, 197] beispielsweise können in ihrer Untersuchung zur Akzeptanz von IS einen starken positiven Zusammenhang zwischen den Konstrukten „Intention to Use“ und „Actual System Usage“ nachweisen. Die substantielle empirische Unterstützung (vgl. auch [Szaj96; ShHa98]) für den starken Zusammenhang zwischen den beiden Konstrukten „Intention to Use“ und „Actual System Usage“ rechtfertigt die beschriebene Schwerpunktsetzung im Rahmen der Akzeptanzforschung. Zur Beschreibung des Konstruktes werden folgende Variablen herangezogen: die Häufigkeit der

Systemnutzung zur Erfüllung konkreter Aufgabenstellungen, die geplante zukünftige Nutzungshäufigkeit des IS sowie die zukünftige Bedeutung des IS zur Grundversorgung mit Informationen.

5 Modellbildung

Die empirischen Untersuchungen zur Akzeptanz von Informationstechnologie stellen die Benutzerfreundlichkeit und die Nützlichkeit eines IS in den Vordergrund der Analyse. Diese Faktoren liefern jedoch einen zu geringen Erklärungsanteil für die Akzeptanz von IS. Das im Rahmen der vorliegenden Arbeit entwickelte Akzeptanzmodell führt daher SYSTEMVERTRAUEN als Modellkomponente ein. Die Aufnahme dieses Faktors basiert auf der Annahme, dass SYSTEMVERTRAUEN einerseits als Zwischenvariable auf die Akzeptanz von IS wirkt, andererseits auch einen direkten Beeinflussungseffekt auf die zukünftige Nutzung des IS ausübt.

Für die erwähnten Beziehungen werden entsprechende Hypothesen formuliert, wobei die Überprüfung auf einem Signifikanzniveau von 5 % erfolgt.

Hypothese 1: Die Benutzerfreundlichkeit eines IS (EASE OF USE) hat einen signifikanten Effekt auf das Vertrauen des Nutzers in das IS (SYSTEMVERTRAUEN).

H₀: Die Benutzerfreundlichkeit eines IS (EASE OF USE) hat keinen signifikanten Effekt auf das Vertrauen des Nutzers in das IS (SYSTEMVERTRAUEN).

Hypothese 2a: Die Benutzerfreundlichkeit eines IS (EASE OF USE) hat einen signifikanten Effekt auf die Nützlichkeit des IS (USEFULNESS).

H₀: Die Benutzerfreundlichkeit eines IS (EASE OF USE) hat keinen signifikanten Effekt auf die Nützlichkeit des IS (USEFULNESS).

Hypothese 2b: Das Vertrauen des Nutzers in ein IS (SYSTEMVERTRAUEN) hat einen signifikanten Effekt auf die Nützlichkeit des IS (USEFULNESS).

H₀: Das Vertrauen des Nutzers in ein IS (SYSTEMVERTRAUEN) hat keinen signifikanten Effekt auf die Nützlichkeit des IS (USEFULNESS).

Hypothese 3a: Die Benutzerfreundlichkeit eines IS (EASE OF USE) hat einen signifikanten Effekt auf die zukünftige Nutzungsabsicht des Systemnutzers (INTENTION TO USE).

H₀: Die Benutzerfreundlichkeit eines IS (EASE OF USE) hat keinen signifikanten Effekt auf die zukünftige Nutzungsabsicht des Systemsnutzers (INTENTION TO USE).

Hypothese 3b: Die Nützlichkeit eines IS (USEFULNESS) hat einen signifikanten Effekt auf die zukünftige Nutzungsabsicht des Systemnutzers (INTENTION TO USE).

H₀: Die Nützlichkeit eines IS (USEFULNESS) hat keinen signifikanten Effekt auf die zukünftige Nutzungsabsicht des Systemnutzers (INTENTION TO USE).

Hypothese 3c: Das Vertrauen des Nutzers in ein IS (SYSTEMVERTRAUEN) hat einen signifikanten Effekt auf die zukünftige Nutzungsabsicht des Systemnutzers (INTENTION TO USE).

H₀: Das Vertrauen des Nutzers in ein IS (SYSTEMVERTRAUEN) hat keinen signifikanten Effekt auf die zukünftige Nutzungsabsicht des Systemnutzers (INTENTION TO USE).

Die formulierten Hypothesen werden in weiterer Folge auf Basis einer empirischen Studie zur Akzeptanz von bankbetrieblichen Informationssystemen getestet.

6 Informationssysteme im bankbetrieblichen Risikomanagement

Der Prozess des bankbetrieblichen Risikomanagements umfasst die Identifikation bzw. Quantifizierung von banktypischen Risiken sowie die Risikosteuerung seitens der Bankzentrale. In der Literatur wird besonders auf die zunehmende Bedeutung risikopolitischer Maßnahmen im Kreditgeschäft hingewiesen: Fischer beispielsweise betont die relevante Stellung des Kreditbereiches im Risikomanagement von Banken [vgl. Fisc00, 109]. Der Einsatz von IS unterstützt die Umsetzung von Maßnahmen im Bereich des Kreditrisikomanagements: insbesondere das Risikoreporting, d.h. die regelmäßige Produktion standardisierter Berichte mit den wesentlichen Kennziffern bzw. der Aufbau von Risikoinformationssystemen mit Branchensorientierung erfüllen wichtige Funktionen in der Bankunternehmensführung. Auch die Beratung der Mitarbeiter aus dem Vertrieb (Privat- und Firmenkundengeschäft) in Fragen des Risikomanagements wird durch die Nutzung von IS erleichtert. Der

Aufgabenschwerpunkt von IS im Kreditrisikomanagement liegt in der systematischen Vermittlung der bankinternen Kreditpolitik. Jede Bank bildet sich ein eigenes Urteil über den erwarteten Verlust bzw. Gewinn mit einem Kreditnehmer, wobei die Kreditpolitik der Bank und die Art und Weise der Kundenbetreuung die Kreditbeziehung entscheidend beeinflussen [vgl. Fisc00, 111 ff.].

Das Risikomanagement in der HVB ist für die eigenständige Steuerung der Risiken im Rahmen der vom Bankvorstand übertragenen Kompetenzen verantwortlich (Portfoliomanagement). Schwerpunkte sind dabei die Entwicklung und Vorgabe geschäftspolitischer Leitlinien (Risikogrundsätze der Kreditpolitik), die laufende Risikoanalyse und die Pflege entsprechender Informationsgrundlagen. Die IS im Kreditrisikomanagement der HVB stützen sich auf strukturierte, steuerungsrelevante Informationen zur Vermittlung der Kreditpolitik der Bank. Auf der Basis von individuell auf Branchen zugeschnittenen Berichten erfolgt die Zuordnung des Kunden zu einer Bonitätsklasse. Die IS zur risikopolitischen Steuerung der Bank werden laufend weiterentwickelt, um eine Verbesserung ihrer Frühwarnfunktionen zu erreichen. Mit der ersten vollständigen Überarbeitung des untersuchten Brancheninformationssystems KoBra im Februar 2000 erfolgte die konzernweite Ausrichtung des Kundenmanagementprozesses nach Branchen in der HVB (zur Branchenorientierung vgl. [UwCh00]). Informationsbeschaffung (Informations- und Dokumentationsstellen) und Informationsverarbeitung (Kreditbearbeitung, Kundenbetreuung) wurden dem Zentralbereich „Risikomanagement“ zugeordnet (zur Informationsverarbeitung vgl. [MoWö99]). Das Brancheninformationssystem KoBra zielt nun darauf ab, den Prozess des Risikomanagements der Bankunternehmensführung zu unterstützen. Von großer Bedeutung ist dabei die Veröffentlichung von verbindlichen Handlungsanweisungen, die den Mitarbeitern die Kreditpolitik der Bank vermitteln.

Die Inhalte der Kobra-Branchenberichte werden von den Mitarbeitern des Brancheninformationszentrums - den Branchenanalysten - laufend gepflegt und aktualisiert. Für die Branchenkapitel „Markt“, „Wettbewerb“, „Rating / Steuerung“ und „Branchenstandards“ sind zusätzlich Branchenmanager bzw. Senior Risk Manager eingebunden, die über weiteres Spezialwissen der Branche verfügen. Dieses Branchenwissen der Senior Risk Manager fließt in die Risikoeinschätzung und Steuerung des Portfolios der Bank ein. Die konzernweite Grundversorgung mit Brancheninformationen erfolgt über 130 Branchenberichte und ca. 2000 Charts und Tabellen. Retrievalmöglichkeiten um-

fassen die Suche über Branchenschlüssel⁵, Volltext- oder Schlagwortsuche sowie die Auswahl nach Gesamtbericht, Charts/Tabellen oder Branchengruppen. Jeder Branchenbericht ist gleich aufgebaut und besitzt folgende - aus fünf Hauptkapiteln bestehende - Informationsstruktur:

- Das Kapitel „Überblick“ dient als Kurzzusammenfassung der Kapitel „Markt“ und „Wettbewerb“. Darüber hinaus werden relevante branchenspezifische Fachbegriffe erläutert.
- Im Kapitel „Markt“ werden zunächst die Größe der jeweiligen Märkte und deren Wachstum bestimmt sowie die Ursachen dieser Entwicklung analysiert. Darauf aufbauend wird eine Prognose der künftigen Entwicklung (Prognosehorizont 5 Jahre) getroffen und wichtige Marktsegmente in der Branche berücksichtigt.
- Ausgangspunkt des Kapitels „Wettbewerb“ sind die Wettbewerbsstrategien einzelner Unternehmen und eine Übersicht der Marktführer. Von besonderem Interesse ist, wie sich diese Strategien im Wettbewerbsergebnis niederschlagen: Auf Basis der Kosten-, Preis- und Gewinnstruktur wird das Insolvenzrisiko der Branche untersucht.
- Das Branchenrating des Kapitels „Rating/Steuerung“ gibt die mittelfristige Risikoeinschätzung einer Branche wieder und wird von den Branchenanalysten unabhängig vom Kreditportfolio der Bank erstellt. Verbindliche Vorgaben für die Kredit-Engagements setzen die Senior Risk Manager aus den Bereichen PuG bzw. FKD mit der „PuG-Branchenampel“ bzw. dem „FKD-Steuerungssignal“. Zusatzinformationen und Ausnahmeregelungen zu den Vorgaben finden sich in allen Branchenberichten. Der Kompetenzträger ist damit verpflichtet, die Kreditvergabe nach den in KoBra veröffentlichten Vorgaben auszurichten.
- Im Rahmen des Kapitels „Branchenstandards“ geben die Senior Risk Manager die Bearbeitungsregeln für den Kreditantrag vor. Diese Branchenstandards sind somit ein verbindlicher Bestandteil der Bonitätsanalyse. Im Mittelpunkt des Kapitels stehen weiters branchentypische Kennzahlen erfolgreicher Unternehmen und Informationen zur Gesamtbeurteilung des Unternehmens.

⁵ Die maximal siebenstelligen Branchenschlüssel entsprechen den EU-konformen WZ-Schlüsseln (Systematik der Wirtschaftszweige). Diese Suchoption wird gewählt, wenn man den Schlüssel kennt oder zumindest weiß, zu welcher Schlüsselgruppe die Information gehört, die benötigt wird.

7 Ergebnisse

Insgesamt haben 542 Mitarbeiter den Fragebogen beantwortet, wobei der Großteil der Respondenten den Bereichen PuG (283 Personen) und FKD (211 Personen) zuzuordnen sind. Die restlichen 48 Antworten betreffen den Immobilienbereich und sonstige Bereiche der Bank. Dies entspricht einer Rücklaufquote von 45,2 %.

Im folgenden Kapitel werden die Ergebnisse der Faktorenanalyse und der Pfadanalyse präsentiert. Auf Basis der Faktorenanalyse erfolgt die Überprüfung der formulierten Hypothesen mittels der Pfadanalyse, die es ermöglicht, kausale Abhängigkeiten zwischen bestimmten Merkmalen zu untersuchen. Der Analyse werden die insgesamt 494 Antworten der Mitarbeiter aus den Bereichen PuG und FKD zugrunde gelegt.

7.1 Faktorenanalyse

Ziel der Faktorenanalyse ist es, aus einer Vielzahl möglicher Variablen die voneinander unabhängigen Einflussfaktoren festzustellen, um die Datenmenge auf wenige wichtige Faktoren reduzieren zu können. Dazu müssen jene Faktoren ermittelt werden, welche die beobachteten Zusammenhänge zwischen den gegebenen Variablen möglichst vollständig erklären. Die Variablen werden dabei jenem Faktor zugeordnet, auf den sie am höchsten laden. Im Rahmen der vorliegenden Analyse wurden Faktorladungen mit einem Wert von größer als 0,4 als relevant angesehen. Als Extraktionsmethode (Eigenwert >1) wurde die Hauptkomponentenanalyse mit orthogonaler Rotation (Varimax-Methode) gewählt. Zur Interpretation der auf einem Faktor hoch ladenden Variablen wird jeweils ein entsprechender Sammelbegriff festgelegt. Die vier ermittelten Konstrukte können inhaltlich als SYSTEMVERTRAUEN, EASE OF USE, USEFULNESS und INTENTION TO USE wiedergegeben werden. Die Ergebnisse der Faktorenanalyse unterstützen mit hohen Faktorladungen die theoretischen Überlegungen zu den Modellkomponenten.

Die Modellkomponente SYSTEMVERTRAUEN wird von Variablen gekennzeichnet, die das Vertrauen des Nutzers in das IS wiedergeben. Die Vorgaben der Steuerungssignale sind ausschließlich in den beiden Branchenskapiteln „Rating/Steuerung“ und „Branchenstandards“ enthalten. Diese Informationen zielen darauf ab, dem Nutzer die konkrete Kreditpolitik der Bank zu vermitteln. Die Faktorladungen der Variablen können wie folgt verstanden werden: Je besser die Handlungsanweisungen bzw. -empfehlungen für den Systemnutzer nachvollziehbar sind, desto eher wird der Nutzer die Kreditvergabe nach

den in den Branchenkapiteln „Rating/Steuerung“ und „Branchenstandards“ veröffentlichten Vorgaben ausrichten.

Der Faktor EASE OF USE repräsentiert die benutzerfreundliche Gestaltung von IS bzw. die Bildung informationeller Mehrwerte. Die zwei am stärksten ausgeprägten Variablen „Struktur der Branchenberichte“ und „Suchmöglichkeiten nach Informationen“ können - der Literatur entsprechend - als „Ease of finding“ interpretiert werden. Die anderen drei Variablen - wie beispielsweise „Arbeitsentlastung bei der Informationsbeschaffung“ - geben die Dimension der informationellen Mehrwertbildung mit Effizienzwirkung wieder.

Der Faktor USEFULNESS wird von Variablen bestimmt, die die Bedeutung der bereitgestellten Informationen zur allgemeinen Entscheidungsunterstützung im Kreditgeschäft wiedergeben. In diesem Zusammenhang zeichnen sich insbesondere die zwei Variablen „Markt“ und „Wettbewerb“ durch hohe Faktorladungen aus. Etwas schwächer ausgeprägt ist die Faktorladung des Branchenkapitels „Überblick“.

Das Konstrukt INTENTION TO USE hängt stark von der Nutzungshäufigkeit vor Kreditanträgen ab. Die Absicht des Nutzers, das IS auch in Zukunft zu verwenden, wird somit geprägt von der Häufigkeit der Systemnutzung vor Kreditanträgen. Variablen wie etwa „zukünftige Nutzungshäufigkeit“ oder „Nutzungshäufigkeit vor Kundenbesuchen“ weisen ebenfalls hohe Faktorladungen aus.

7.2 Pfadanalyse

Die Pfadanalyse ist ein multivariates statistisches Verfahren zur Überprüfung von kausalen Abhängigkeiten zwischen bestimmten Merkmalen. Das hypothetische Kausalmodell beruht dabei auf fundierten theoretischen Überlegungen. Die Beziehungen zwischen den Konstrukten werden in Folge durch die Berechnung von standardisierten Koeffizienten (Beta-Koeffizienten) und Varianzen ermittelt.

Hypothese 1:

Das Konstrukt EASE OF USE hat einen signifikanten Effekt auf SYSTEMVERTRAUEN ($p < 0,001$ und $\text{Beta} = 0,44$). Der Pfad erklärt 19 % der Varianz in SYSTEMVERTRAUEN.

⇒ Die Nullhypothese, nach der die Benutzerfreundlichkeit eines IS (EASE OF USE) keinen signifikanten Effekt auf das Vertrauen des Nutzers in das IS (SYSTEMVERTRAUEN) hat, ist zu verwerfen.

Hypothesen 2a und 2b:

Die Konstrukte EASE OF USE und SYSTEMVERTRAUEN haben einen signifikanten Effekt auf USEFULNESS ($p < 0,001$). Die entsprechenden Pfadkoeffizienten betragen 0,48 und 0,25. Die erklärte Varianz beträgt 37 %.

Hypothese 2a:

⇒ Die Nullhypothese, nach der die Benutzerfreundlichkeit (EASE OF USE) keinen signifikanten Effekt auf die Nützlichkeit des IS (USEFULNESS) hat, ist zu verwerfen.

Hypothese 2b:

⇒ Die Nullhypothese, nach der das Vertrauen des Nutzers in das IS (SYSTEMVERTRAUEN) keinen signifikanten Effekt auf die Nützlichkeit des IS (USEFULNESS) hat, ist zu verwerfen.

Hypothesen 3a, 3b und 3c:

Der Beeinflussungseffekt der drei Konstrukte EASE OF USE, USEFULNESS und SYSTEMVERTRAUEN auf INTENTION TO USE ist signifikant. EASE OF USE und USEFULNESS erzielen beinahe identische Werte ($p < 0,01$ und $\text{Beta} = 0,20$) hinsichtlich des direkten Effektes auf INTENTION TO USE. Etwas schwächer ausgeprägt ist die Korrelation zwischen SYSTEMVERTRAUEN und INTENTION TO USE ($\text{Beta} = 0,13$ bei $p < 0,05$). Die durch die drei Pfade erklärte Varianz beträgt 20 %. Der Anteil der erklärten Varianz ist bei anderen Untersuchungen im Bereich TAM weniger stark ausgeprägt: Lederer et al. beispielsweise erzielen in ihrer Untersuchung einen Wert von 15 % [vgl. LeMa00, 275].

Hypothese 3a:

⇒ Die Nullhypothese, nach der die Benutzerfreundlichkeit eines IS (EASE OF USE) keinen signifikanten Effekt auf die zukünftige Nutzungsabsicht des Systemanwenders (INTENTION TO USE) hat, ist zu verwerfen.

Hypothese 3b:

⇒ Die Nullhypothese, nach der die Nützlichkeit eines IS (USEFULNESS) keinen signifikanten Effekt auf die zukünftige Nutzungsabsicht des Systemanwenders (INTENTION TO USE) hat, ist zu verwerfen.

Hypothese 3c:

⇒ Die Nullhypothese, nach der das Vertrauen des Nutzers in ein IS (SYSTEMVERTRAUEN) keinen signifikanten Effekt auf die zukünftige Nutzungsabsicht des Systemanwenders (INTENTION TO USE) hat, ist zu verwerfen.

Ein Vergleich der Korrelationen zwischen den Parametern des Modells zeigt, dass die Beziehungen zwischen EASE OF USE und SYSTEMVERTRAUEN bzw. zwischen SYSTEMVERTRAUEN und USEFULNESS einen ähnlich starken Zusammenhang aufweisen. Damit wird die Bedeutung von SYSTEMVERTRAUEN als Zwischenvariable zur Erklärung des Akzeptanzmodells deutlich.

8 Resümee

Durch eine kritikgeleitete Verknüpfung von bisherigen Forschungsergebnissen mit einer Erweiterung für den Bereich der Managementinformationssysteme um den Faktor SYSTEMVERTRAUEN wird ein neu entwickeltes Akzeptanzmodell erarbeitet und anhand einer repräsentativen empirischen Studie im Bankensektor überprüft.

Den Schwerpunkt der Arbeit bildet die Überprüfung von sechs Hypothesen, wobei die postulierten Beziehungen zwischen den einzelnen Konstrukten des Akzeptanzmodells anhand multivariater Auswertungen untersucht werden. Dazu werden in einem ersten Schritt die neu eingeführte Einflussgröße SYSTEMVERTRAUEN sowie die Konstrukte EASE OF USE, USEFULNESS und INTENTION TO USE theoretisch erarbeitet und miteinander in Verbindung gesetzt. Die empirische Validierung der Konstrukte erfolgt mittels Faktorenanalyse, deren Resultate die theoretischen Überlegungen zu den Modellkomponenten durch hohe Faktorladungen unterstützen. Im Rahmen der Kernhypothesen wird überprüft, ob die einzelnen Modellkomponenten signifikante Beeinflussungseffekte untereinander ausüben. Mit Hilfe der Pfadanalyse wird die Erklärungsqualität des Akzeptanzmodells untersucht und der Zusammenhang zwischen den Faktoren aufgezeigt. Gemäß Hypothese 1 wird ein signifikanter Effekt des Konstruktes EASE OF USE auf SYSTEMVERTRAUEN statistisch nachgewiesen. Hinsichtlich der Hypothesen 2a und 2b werden die signifikanten Beeinflussungseffekte der Faktoren EASE OF USE und SYSTEMVERTRAUEN auf USEFULNESS aufgezeigt. Schließlich führen auch die Beeinflussungseffekte der drei Komponenten EASE OF USE, USEFULNESS und SYSTEMVERTRAUEN auf die abhängige Variable INTENTION TO USE (Hypothesen 3a, 3b und 3c) zu signifikanten Resultaten. Eine Analyse der direkten und indirekten Effekte zwischen den Konstrukten führt zu dem Ergebnis, dass die

Faktoren EASE OF USE und SYSTEMVERTRAUEN die zentralen Bestimmungsgrößen des entwickelten Akzeptanzmodells sind. Ein Vergleich der Pfadkoeffizienten macht die Bedeutung des neu eingeführten Faktors SYSTEMVERTRAUEN zur Erklärung des Akzeptanzmodells deutlich. Die positiven und stark ausgeprägten Beziehungen zwischen EASE OF USE und SYSTEMVERTRAUEN bzw. SYSTEMVERTRAUEN und USEFULNESS bestätigen die erfolgreiche Erweiterung des Modells. Die multivariate Analyse zeigt, dass SYSTEMVERTRAUEN einerseits als wichtige Zwischenvariable auf die Akzeptanz von Informationssystemen wirkt, andererseits auch einen direkten Beeinflussungseffekt auf die zukünftige Nutzung des Informationssystems ausübt.

Die Studie hat zu unmittelbaren Konsequenzen in der Praxis geführt: Die entsprechende Abteilung in der Bayerischen HypoVereinsbank wurde personell beträchtlich aufgestockt und die Ergebnisse direkt zu einer Systemverbesserung herangezogen. Die vollständige Neuausrichtung der System-unterstützten Branchenorientierung beinhaltet u.a. die differenziertere Darstellung von Märkten in jenen Branchen in denen der Bankkonzern überdurchschnittlich wachsen will (Zielbranchen), die klare Orientierung an jene Branchen, in denen die Bank einen Portfolioschwerpunkt hat (Kernbranchen), die Erhöhung der Analyseintensität der Branchenberichte sowie die verstärkte Individualanalyse bei größeren Kreditengagements. Die signifikanten Resultate der empirischen Studie haben damit insgesamt eine unmittelbare Systemverbesserung erzielt.

9 Literatur

- [Acko67] Ackoff R.L.: Management Misinformation Systems, in: Management Science, 14, 4, 1967, 147 - 156.
- [Alex98] Alex B.: Künstliche neuronale Netze in Management-Informationssystemen. Wiesbaden: Gabler, 1998.
- [Alte99] Alter S.: Information systems: a management perspective. Reading: Addison Wesley Longman, 1999.
- [Back96] Backhaus K. et al.: Multivariate Analysemethoden. Berlin: Springer, 1996.
- [DaBa89] Davis F.D./Bagozzi R.P./Warshaw P.R.: User acceptance of computer technology: A comparison of two theoretical models, in: Management Science, 35, 8, 1989, 982 - 1003.
- [Davi86] Davis F.D.: A technology acceptance model for empirically testing new end-user information systems: theory and results. Doctoral Dissertation, Sloan School of Management, MA, 1986.
- [Davi93] Davis F.D.: User acceptance of information technology: system characteristics, user perceptions and behavioral impacts, in: Int.J.Man-Machine Studies, 38, 1993, 475-487.

- [DeMc92] DeLone W./McLean E.: Information systems success: The Quest for the dependent variable, in: Information Systems Research, 3, 1, 1992, 60-95.
- [DyCh00] Dyer J.H./Chu W.: The determinants of trust in supplier-automaker relationships in the U.S., Japan, and Korea, in: Journal of International Business Studies, 31, 2, 2000, 259 - 285.
- [FiAj75] Fishbein M./Ajzen I.: Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research. Reading: Addison-Welsey, 1975.
- [Fisc00] Fischer T.R.: Finanzindustrie und Risikomanagement, in: ÖBA, 2, 2000, S 107 - 114.
- [Gati94] Gatian A.: Is user satisfaction a valid measure of system effectiveness? In: Information & Management, 26, 1994, 119 - 131.
- [GoRe00] Govindarajulu C./Reithel B.J./Sethi V.: A model of end user attitudes and intentions toward alternative sources of support, in: Information & Management, 37, 2000, 77 - 86.
- [Hilb84] Hilbig W.: Akzeptanzforschung neuer Bürotechnologien, in: Office Management, 4, 1984, 320 - 323.
- [Kief00] Kiefer T.: Electronic Commerce - Vertrauen als Erfolgsfaktor, in: Die Bank, 5, 2000, 308 - 312.
- [Koll00] Kollmann T.: Die Messung der Akzeptanz bei Telekommunikationssystemen, in: JFB, 2, 2000, 68 - 78.
- [KoWi97] Kortzfleisch H./Winand U.: Kooperieren und Lernen im Intranet, in: Information Management, 2, 1997, 28 - 35.
- [Kuhl89] Kuhlen R.: Pragmatischer Mehrwert von Information. Sprachspiele mit informationswissenschaftlichen Grundbegriffen. Report. Universität Konstanz, Fachbereich Informationswissenschaft, 1989.
- [Kuhl95] Kuhlen R.: Informationsmarkt: Chancen und Risiken der Kommerzialisierung von Wissen. Konstanz: Univ.Verlag, 1995.
- [Kuhl99] Kuhlen R.: Die Konsequenzen von Informationsassistenten. Frankfurt am Main: Suhrkamp, 1999.
- [LeMa00] Lederer A./Maupin D.J./Sena M.P./Zhuang Y.: The technology acceptance model and the World Wide Web, in: Decision Support System, 29, 2000, 269 - 282.
- [LiLu00] Lin J.C.C./Lu H.: Towards an understanding of the behavioural intention to use a web site, in: International Journal of Information Management, 2000, 20, 3, 197 - 298.
- [Luhm88] Luhmann N.: Familiarity, confidence, trust: Problems and alternatives, in: Gambetta (Hg.): Trust: Making and breaking cooperative relations, 1988, 94 - 107.
- [MaSk98] Macy M.W./Skvoretz J.: The evolution of trust and cooperation between strangers: A computational model, in: American Sociological Review, 63, 1998, 638 - 660.
- [MaSn89] Mahmood M.A./Snizek J.: Defining decision support systems: an empirical assessment of end-user satisfaction, in: Information Systems and Operational Research, 27, 3, 1989, 253 - 271.
- [Math91] Mathieson K.: Predicting User Intentions : Comparing the technology acceptance model with the theory of planned behavior“, in: Information Systems Research, 2, 3, 1991, 173 - 191.

- [Meye99] Meyer 1999 Meyer J.-A.: Visualisierung von Informationen: Verhaltenswissenschaftliche Grundregeln für das Management. Wiesbaden: Gabler 1999.
- [MoWö99] Moormann J./Wölfling D.: Auf dem Weg zur neuen Informatik, in: Die Bank, 7, 1999, 462 - 467.
- [Porte85] Porter M.: Competitive Advantage: Creating and Sustaining Superior Performance. New York: The Free Press, 1985.
- [Ratn98] Ratnasingham P.: The importance of trust in electronic commerce, in: Internet Research: Electronic Networking Applications and Policy, 8, 4, 1998, 313 - 321.
- [Reag96] Reagle J. M.: Trust in electronic markets. In: first monday, 2, 1996, <http://www.firstmonday.org/issues/issue2/markets/index.html> (Version 10.1.00).
- [RoHe00] Roberts P./Hendersen R.: Information technology acceptance in a sample of government employees: a test of the technology acceptance model, in: Interacting with computers, 12, 2000, 427 - 443.
- [Schm92] Schmidt R.: Modelle der Informationsvermittlung. Analyse und Bewertung eines experimentellen Programms. Physica-Verlag: Heidelberg, 1992.
- [ShaWe49] Shannon C.E./Weaver W.: The mathematical theory of communication. University of Illinois, Urbana: Ill, 1949.
- [ShHa98] Sheppard B.H./Hartwick J./Warshaw P.R.: The theory of reasoned action: A meta-analysis of past research with recommendations for modifications and future research, in: Journal of Consumer Research, 15, 1998, 325 - 343.
- [Stoc99] Stock W.: Vertrauensmanagement auf Informationsmärkten, in: password, 12, 1999, 21 - 27.
- [Szaj96] Szajna B.: Empirical Evaluation of the Revised Technology Acceptance Model, in: Management Science, 42, 1, 1996, 85 - 92.
- [UwCh00] Uwe C.: Wettbewerbsvorteile durch Markt und Branchenkompetenz. In: Die Bank, 2000, 2, 134 - 137.
- [VeDa00] Venkatesh V./Davis F.D.: A theoretical extension of the technology acceptance model: four longitudinal field studies, in: Management Science, 46, 2, 2000, 186 - 203.
- [Wers74] Wersig G.: Information - Kommunikation - Dokumentation. Verlag Dokumentation: Pullach, 1974.
- [WiPo98] Winand U./Pohl W.: Die Vertrauensproblematik in elektronischen Netzwerken, in: Link J.(Hg.): Wettbewerbsvorteile durch Online-Marketing, Berlin: Springer, 1998.
- [YeJo95] Ye L.R./Johnson P.E.: The impact of explanation facilities on user acceptance of expert systems advice, in: MIS Quarterly, June 1995, 157 - 172.



Informations- und Informationstechnologie- Management in österreichischen Banken, Versicherungen und KFZ-Unternehmen

Christian Schlögl

Karl-Franzens-Universität Graz
Institut für Informationswissenschaft
Universitätsstraße 15, F3
A-8010 Graz
christian.schloegl@uni-graz.at

Zusammenfassung

In diesem Beitrag werden Teile einer empirischen Studie vorgestellt, die das Informationsmanagement (IM) in österreichischen Großunternehmen zum Gegenstand hatte. Die Studie basierte auf einem umfassenden IM-Verständnis. Demzufolge wurden sowohl Informationsvermittlungsstellen als auch EDV-/Organisationsabteilungen auf die Umsetzung einer Reihe von IM-Indikatoren - strategische Informations(technologie)planung, Informationsrichtlinien, Informations(verarbeitungs)controlling, methodische Unterstützung und organisatorische Aspekte - untersucht, die zuvor auf Grundlage eines umfassenden Literaturstudiums herausgearbeitet wurden. Die Ergebnisse zeigen die starke Betonung der technologischen Aspekte im Rahmen des Informationsmanagements auf. Demnach werden die einzelnen IM-Kriterien von den EDV-Abteilungen in einem hohen Maße erfüllt. Hingegen ist bei den Informationsvermittlungsstellen oft schon die Lokalisierung mit Schwierigkeiten verbunden. In einigen Unternehmen konnte nicht einmal eine Stelle eruiert werden.

1 Einleitung

In der Literatur liegen mittlerweile zahlreiche konzeptionelle Beiträge zum Informationsmanagement (IM) vor. Im Vergleich dazu ist die Anzahl empirischer Belege relativ gering. Diese behandeln zudem oft nur technologische Aspekte des Informationsmanagements. Es gibt aber nach wie vor nur wenige



Studien, denen ein umfassendes IM-Verständnis zugrunde liegt. Dies war der Ausgangspunkt für die hier vorgestellte empirische Untersuchung. Ziel dieser Studie war es, den Wissensstand zum Informationsmanagement im Sinne eines umfassenden Managements von Information *und* Informationstechnologie zu erhöhen. Dabei sollten durch die interdisziplinäre Betrachtung des Informationsmanagements sowohl die technologieorientiertere Sichtweise der Wirtschaftsinformatik als auch die informationsorientiertere Perspektive der Informationswissenschaft berücksichtigt werden. Unter anderem wurden folgende Aspekte des Informationsmanagements untersucht: Selbstverständnis, Ausprägungsgrad, Rolle des Endbenutzers, externe Einflussfaktoren, Problembereiche und zukünftige Entwicklung. Aus Platzgründen werden in diesem Aufsatz nur jene Teile der Studie vorgestellt, die den Ausprägungsgrad des Informationsmanagements zum Gegenstand hatten.¹

2 Konzeptualisierung und Untersuchungsplan

Grundlage einer jeden wissenschaftlichen Arbeit ist ein klares Begriffsverständnis. Diese Anforderung gilt insbesondere für das Informationsmanagement, da es gerade hier ein hohes Maß an semantischen „Unstimmigkeiten“ gibt. Aus diesem Grund wurde bei der Konzeptualisierung der Studie einer klaren Begriffsfestlegung und der Operationalisierung der Grundbegriffe große Aufmerksamkeit geschenkt.

2.1 Objekte des Informationsmanagements

Die Studie ging nicht zuletzt aufgrund der Ergebnisse einer szientometrischen Analyse zum Informationsmanagement² von einem umfassenden IM-Verständnis aus. Demnach umfasst Informationsmanagement das Leitungshandeln in einer Organisation in Bezug auf die Bereitstellung und Nutzung der „*informationseinhaltlichen*“ (Datenbanken, Dokumentationen, ...) und „*informationstechnologischen*“ (Hardware, Software, ...) Infrastruktur und der darauf basierenden Informationssysteme. Aus dieser Definition leiten sich Information und Informationstechnologie als die beiden zentralen Objekte des Informationsmanagements ab.

¹ Eine vollständige Darstellung der Studie findet sich bei Schlögl 2001. Alle in diesem Aufsatz enthaltenen Tabellen wurden dieser Publikation entnommen. In einer weiteren Publikation (Schlögl 2002) wurden nur die Ergebnisse der Informationsvermittlungstellen vorgestellt.

² Vgl. Schlögl 2000.

Bei der empirischen Untersuchung wurde dem dadurch entsprochen, dass in diese alle Stellen/Abteilungen einbezogen wurden, die mit dem Management von Information *und* Informationstechnologie betraut sind. Im weiteren Verlauf dieses Beitrags wird dafür die Bezeichnung „*Informationsfunktion*“ verwendet. Es handelt sich dabei um Stellen/Abteilungen, die für andere Abteilungen (gegebenenfalls auch externe Kunden) „*Informationsdienstleistungen*“ erbringen.³ Je nach Ausrichtung (Information oder Informationstechnologie) kann zwischen informations- und technologieorientierten Informationsfunktionen unterschieden werden. Erstere sind dafür zuständig, Informationen primär für andere Stellen im Unternehmen zu beschaffen, zu analysieren, zu speichern, bedarfsgerecht aufzubereiten und weiterzuvermitteln und/oder die dafür erforderliche Infrastruktur aufzubauen und zu betreuen. Ein Beispiel für eine informationsorientierte Informationsfunktion ist die Unternehmensbibliothek, eine Dokumentationsabteilung oder eine Stelle Wissensmanagement. Bei Letzteren handelt es sich typischerweise um die EDV-Abteilung, die für Bereitstellung und Betrieb von computerbasierten Informationssystemen verantwortlich ist. Falls es eine eigene Organisationsabteilung gab, wurde diese ebenfalls zu den *technologieorientierten Informationsfunktionen* gezählt.

2.2 Operationalisierung

In Anlehnung⁴ an Bergeron und Horton wurden zur Beurteilung des Ausprägungsgrads des Informationsmanagements folgende Indikatoren verwendet:

- strategische Informations(technologie)planung,
- Informationsrichtlinien,
- Informations(verarbeitungs)controlling,
- methodische Unterstützung und
- aufbauorganisatorische Aspekte.

EDV-/Organisationsabteilungen wurden darüber hinaus auf nur sie betreffende Kriterien (IT-Management) untersucht.

³ Ein Mitarbeiter der Forschungsabteilung beispielsweise, der neben seiner herkömmlichen Tätigkeit Recherchen in einer Patentdatenbank durchführt, wurde laut dieser Festlegung nicht als Informationsfunktion betrachtet und war damit nicht Gegenstand dieser Untersuchung.

⁴ Die hier gewählten Indikatoren entsprechen jenen von Bergeron und Horton nicht vollständig. Bergeron verwendete in ihrer Studie „Planung“, „Informationsrichtlinien“, „Kosten und Nutzen“, „Organisation“ sowie „Informationsbedarf und –verwendung“ (vgl. Bergeron 1996, S. 107). Laut Horton kann ein IM-Ansatz die Komponenten „Planung“, „Informationsrichtlinien“, „Informations-Kostenartenrechnung und Budgetierung“.

Strategische Informations(technologie)planung:

Laut Heinrich und Lehner umfasst die strategische (Gesamt)Planung die Formulierung der strategischen Ziele, die Strategie-Entwicklung und die strategische Maßnahmenplanung. Letztere wird im Zuge der weiteren Planung schrittweise verfeinert und bis auf die operative Ebene hinunter gebrochen.⁵ Eine wichtige Teilaufgabe der strategischen Maßnahmenplanung ist die Planung der Informationsressourcen (Hardware, Anwendungen, Daten/Information, Personal, etc.).⁶ Ein Hilfsmittel dazu stellen Verzeichnisse (zum Beispiel ein Data-Dictionary) dar, mit denen die Informationsressourcen verwaltet werden.

Eine wichtige Forderung an die strategische Informationsplanung ist deren Ausrichtung an der Unternehmensplanung.⁷ Ein weiterer Aspekt besteht in der Wettbewerbsorientierung.⁸ Demnach sollen durch den Einsatz von Information/Informationstechnologie Wettbewerbsvorteile erreicht, zumindest aber eine Verschlechterung der eigenen Wettbewerbssituation verhindert werden.⁹

Der Status-quo der strategischen Informations(technologie)planung wurde anhand folgender Fragen erhoben:

- Gibt es ein Verzeichnis der Informationsressourcen? Wird ein bestimmtes Werkzeug verwendet?
- Existiert ein schriftlich ausformulierter Informations(technologie)plan?
- Falls ja:
- Welche Arten von Informations(technologie)plänen gibt es: operative und/oder strategische?
- Basiert die Informations(technologie)planung auf der Unternehmensstrategie und den Unternehmenszielen?
- Arbeitet der Leiter der Informationsfunktion direkt bei der Unternehmensplanung mit?
- Wird im Unternehmen aktiv versucht, durch den Einsatz von Information und/oder Informationstechnologie Wettbewerbsvorteile zu erreichen? Falls ja: in welchen Bereichen?

⁵ Vgl. Heinrich 1999; Lehner 1993, S. 57 ff.

⁶ Vgl. Heinrich 1999, S. 128.

⁷ Vgl. zum Beispiel Bergeron 1996, S. 276 f.; Heinrich 1999, S. 118; Lehner 1992, S. 97; Marchand / Horton 1986, S. 190 ff.; Wigand 1988, S. 27.

⁸ Synnott sieht zwischen der Integration von Informations- und Unternehmensplanung und dem wettbewerbsorientierten Einsatz von Information/Informationstechnologie eine Mittel-/Zweckbeziehung. Erstere ist Voraussetzung dafür, dass Wettbewerbsvorteile lukriert werden können (vgl. Synnott 1987, S. 11).

⁹ Vgl. Lehner 1993, S. 22.

Informationsrichtlinien:

Informationsrichtlinien sind ein weiterer wichtiger Bestandteil eines IM-Konzepts.¹⁰ Es handelt sich dabei um in einer Organisation gültige Grundsätze für den Bereich der Informationsverarbeitung. Der Hauptzweck von Informationsrichtlinien besteht einerseits darin, den Endbenutzern eine Orientierung zu geben, was sie von der Informationsverarbeitung erwarten können bzw. was von ihnen erwartet wird. Andererseits versorgen sie das Informationsmanagement mit Normen, an denen es sein Handeln ausrichten kann.

Bei der empirischen Untersuchung wurden folgende Aspekte berücksichtigt:

- Vorhandensein von Informationsrichtlinien,
- Inhalte/Bereiche,
- Bekanntheitsgrad und
- Bindung.

Informations(verarbeitungs)controlling:

Viele Autoren sind der Meinung, dass Informationen Kosten verursachen und einen Wert haben.¹¹ Zur Umsetzung dieses Grundsatzes bedarf es allerdings eines geeigneten Instrumentariums. Das „herkömmliche“ Rechnungswesen ist kaum geeignet, Informationsdienste und –produkte ausreichend zu berücksichtigen. Die Bestimmung der Informationskosten ist aber eine wichtige Voraussetzung für ein Informationscontrolling, dessen Entwicklungsstand sich laut Spitta durch die Variablen Leistungsverrechnung und Basiskennzahlen feststellen lässt.¹²

Zur Analyse des Informations(verarbeitungs)controlling wurden folgende Indikatoren verwendet:

- Form, in der die jeweilige Informationsfunktion geführt wird (Cost Center oder Profit Center),
- Höhe des Budgets,
- Planung und Steuerung der (Informations)Kosten,
- Ermittlung von Kennzahlen und
- Kostenverrechnung.

¹⁰ Vgl. zum Beispiel Horton 1985, S. 227; Lytle 1988, S. 11; Orna 1990.

¹¹ Vgl. zum Beispiel Horton 1985, S. 18; Lytle 1986, S. 391; Marchand/Horton 1986, S. 206 ff.; Taylor 1986, S. 175.

¹² Vgl. Spitta 1998, S. 426.

Methoden:

Von Seiten der Informationswissenschaft ist das Methodenangebot für das Informationsmanagement relativ gering. Hingegen werden im Bereich des technologieorientierten Informationsmanagements mittlerweile zahlreiche Methoden bereitgestellt.¹³ Bei der empirischen Untersuchung wurden diese in folgende Gruppen eingeteilt:

- strategische Informations(system)planung,
- Informations(verarbeitungs)controlling,
- allgemeine und sonstige Methoden.

Organisation:

Damit das Informationsmanagement vor allem in größeren Organisationen seinen umfassenden Aufgaben nachkommen kann, ist eine strukturorganisatorische Verankerung erforderlich. Allgemein gilt, dass die Stellung einer Organisationseinheit davon abhängt, für wie wichtig ihr Beitrag für das Unternehmen empfunden wird. Eine hohe organisatorische Verankerung des Informationsmanagements deutet daher auf eine dementsprechende Bedeutung für das jeweilige Unternehmen hin.

Bei der empirischen Studie wurden folgende aufbauorganisatorischen Aspekte untersucht:

- Wie viele und welche Stellen/Abteilungen befassen sich mit Aufgaben des Informationsmanagements?
- Wie viele Mitarbeiter haben diese Stellen/Abteilungen?
- Wie sind sie strukturorganisatorisch in das Unternehmen eingegliedert?
- Wurden einzelne Aufgabenbereiche oder ganze Informationsfunktionen in der Vergangenheit ausgelagert?

IT-Management:

Unter dieser Bezeichnung wurden mehrere Kriterien zusammengefasst, die ausschließlich EDV- und Organisationsabteilungen betreffen. Es sind dies:

- unternehmensweite IT-Standards und deren Einhaltung,
- Informations(system)architektur,
- Stelle/Abteilung Datenmanagement sowie
- Monitoring der Technologieentwicklung.

¹³ Ein umfangreicher Überblick über das Methodenangebot findet sich bei Heinrich (1999, S. 327 ff.), Lehner (1993, S. 101 ff.) und Biethahn/Muksch/Ruf (1996, S. 235 ff.).

Alle diese Indikatoren haben ihre Entsprechung beim IM-Konstrukt von Lewis et al.¹⁴ Unternehmensweite Standards für den Einsatz von Informationstechnologie sind für ein reibungsloses Zusammenwirken der einzelnen Komponenten der Informationsinfrastruktur besonders wichtig. Die Informationsarchitektur wurde nicht nur in der Literatur besonders thematisiert, sondern auch bei einigen empirischen Studien¹⁵ als ein Kernpunkt des Informationsmanagements identifiziert. Ein wichtiger Teilbereich des Informationsmanagements ist das Datenmanagement.¹⁶ Dieser Bedeutung sollte dadurch Rechnung getragen werden, dass für die im Rahmen des Datenmanagements anfallenden Aufgaben eine eigene Stelle/Subabteilung geschaffen wird.¹⁷ Das Beobachten der Entwicklung auf dem Technologiemarkt ist laut Heinrich eine Teilaufgabe des strategischen Informationsmanagements. Ziel ist es, neue Entwicklungen auf dem IT-Markt bereits anhand schwacher Signale, zumindest aber nicht später als die Mitbewerber zu erkennen, und die sich bietenden Chancen für das eigene Unternehmen abzuschätzen.¹⁸

Das Ergebnis der Operationalisierung ist in Abbildung 1 dargestellt.

¹⁴ Vgl. Lewis/Snyder/Rainer 1995, S. 210 ff.

¹⁵ Vgl. zum Beispiel Brancheau/Wetherbe 1987; Niederman/Brancheau/Wetherbe 1991.

¹⁶ Vgl. zum Beispiel Heinrich 1999.

¹⁷ Vgl. Gemünden/Schmitt 1991, S. 28; Niedereichholz/Wentzel 1985, S. 285 ff.

¹⁸ Vgl. Heinrich 1999, S. 156 f.

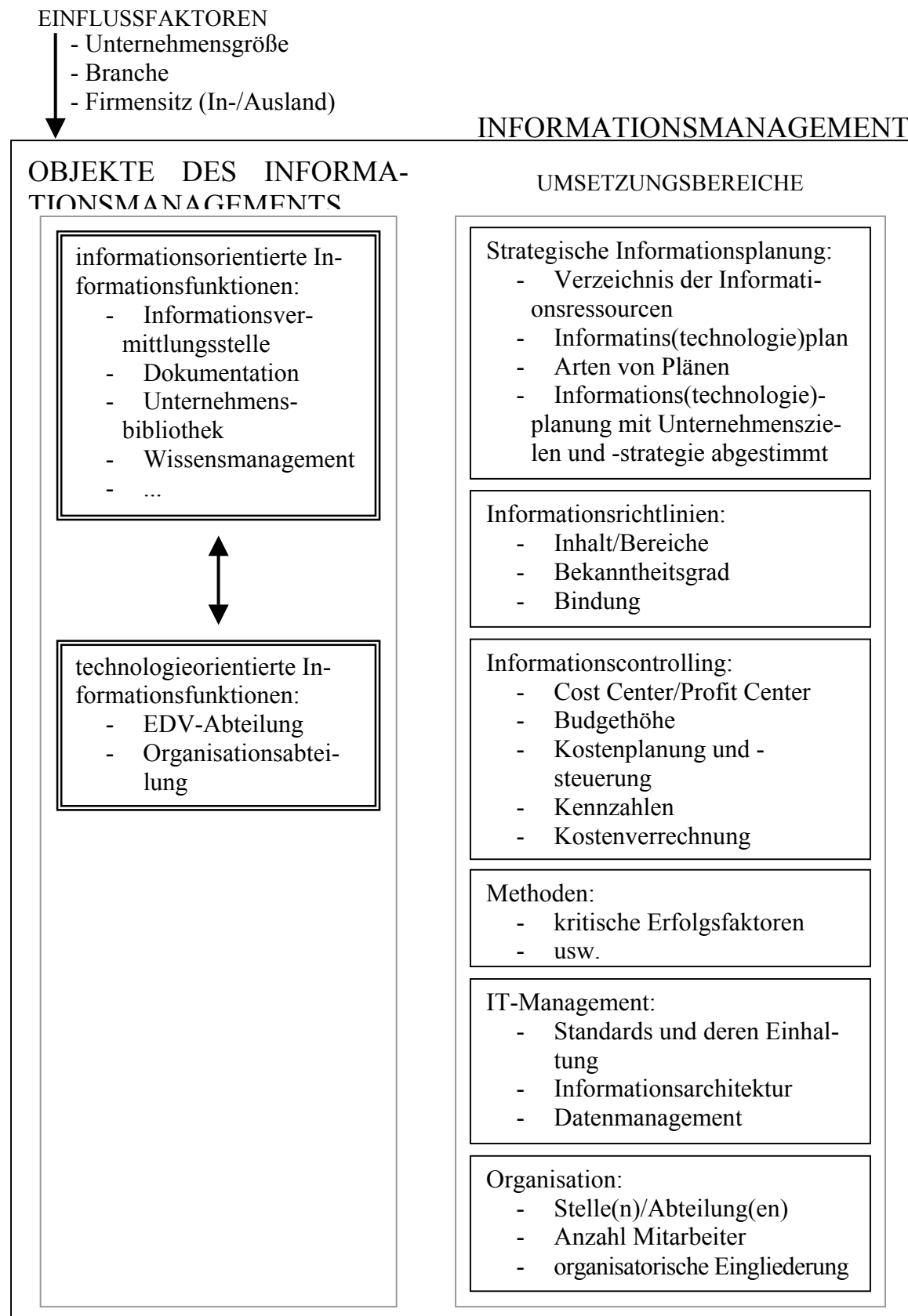


Abbildung 1: Modell eines umfassenden Informationsmanagements (in Anlehnung an Schlögl 2001)

2.3 Untersuchungsplan

Da die Kenntnisse zum Untersuchungsgegenstand nicht ausreichend waren, erschien eine analytische Studie nicht sinnvoll. Auch der Umstand, dass das Identifizieren der informationsorientierten Informationsfunktionen mit Schwierigkeiten verbunden sein könnte¹⁹ und dadurch ein großes „sample“, wenn überhaupt, dann nur mit einem sehr hohen Aufwand realisierbar gewesen wäre, sprach gegen eine hypothesenprüfende Studie. Da zu erwarten war, dass auch in der Praxis kein einheitliches Begriffsverständnis besteht, kam eine schriftliche Befragung nicht in Frage. Zur Datenerhebung – diese erfolgte im Frühjahr 2000 – wurden strukturierte und großteils standardisierte Interviews eingesetzt. Ein Vorteil dieser Erhebungsmethode bestand auch darin, dass beim (bei den) ersten Interview(s) in einem Unternehmen allenfalls weitere noch nicht identifizierte Informationsfunktionen eruiert werden konnten.

	Banken	Versicherungen	KFZ	Σ
Anzahl Unternehmen	6	7	5	18
Informationsfunktionen (IF)	19	14	9	42
technologieorientierte IF	6	9	5	20
informationsorientierte IF	13	5	4	22
EDV ausgelagert	3	4	2	9

Tabelle 1: Informationsfunktionen – branchenweiser Vergleich

Bei der Bestimmung der Grundgesamtheit wurde davon ausgegangen, dass ein institutionalisiertes Informationsmanagement nur in größeren Unternehmen vorhanden ist. Da auch ein sektoraler Vergleich durchgeführt werden sollte, wurde die Grundgesamtheit folgendermaßen definiert: *alle Banken, (Privat)Versicherungen und Unternehmen der Automobilindustrie, die in Österreich mehr als 1000 Mitarbeiter beschäftigen*. Die zu untersuchenden Unternehmen wurden durch eine Datenbankrecherche in der Internetversion der KSV-Datenbank²⁰ ermittelt. Die Recherche ergab insgesamt 19 Unternehmen, wovon nur ein Unternehmen der KFZ-Branche die Mitarbeit an der Studie ablehnte. Die teilnehmenden Unternehmen setzten sich aus sechs Banken, sieben Versicherungen und fünf Unternehmen der Automobilindustrie zusammen. Die durchschnittliche Beschäftigtenzahl lag in den Banken bei über 4.200, in den Versicherungen bei ca. 3.400 und in den KFZ-Unternehmen bei ungefähr 2.900 Mitarbeiter. Insgesamt haben sich 42 Personen, großteils die Leiter der Informationsfunktionen, zur Mitarbeit an der Untersuchung bereit-erklärt, nur sechs haben ihre Mitwirkung abgelehnt. Das entspricht einer sehr

¹⁹ Vgl. Bergeron 1996, S. 9; Lewis/Martin 1989; Schlögl/Voglmayr 1999.

²⁰ <http://marketing.ksv.at/search.asp?Action=NeueAbfrage> (Stand: März 2000).

hohen „Beteiligungsquote“ von 88 %. 20 der 42 Interviewpartner kamen von technologieorientierten, 22 von informationsorientierten Informationsfunktionen.

3 Ergebnisse

Nachfolgend werden die Resultate zum Ausprägungsgrad des Informationsmanagements präsentiert. Bei der Darstellung der einzelnen Umsetzungsbereiche erfolgt mit Ausnahme der organisatorischen Aspekte eine Trennung in informations- und technologieorientierten Informationsfunktionen, wobei am Beginn jeweils ein Ergebnisüberblick in Tabellenform gegeben wird.²¹

3.1 Organisatorische Aspekte

Ein erstes interessantes Ergebnis besteht darin, dass die *EDV-Abteilung* in genau der Hälfte der Unternehmen *ausgelagert* ist. Die Interviews ergaben aber, dass eine ausgelagerte EDV normalerweise wie eine herkömmliche IT-Abteilung betrachtet wird. Im Vergleich zu den technologieorientierten Informationsfunktionen vermitteln die *informationsorientierten* ein sehr unterschiedliches Bild. So sind von den 22 informationsorientierten Informationsfunktionen nur zwei Stellen *ausgelagert*. Der Grund für die niedrige „Outsourcing-Quote“ dürfte darin liegen, dass es grundsätzlich weniger problematisch ist, das Management der Informationstechnologie außer Haus zu geben. Hingegen ist dies bei Information als entscheidenden Erfolgsfaktor für ein Unternehmen nur beschränkt möglich.

Überraschend ist, dass „traditionelle“ Informationsvermittlungsstellen in den untersuchten Unternehmen kaum zu finden sind. In nur zwei Unternehmen gibt es eine (zentrale) Unternehmensbibliothek. Neben der reinen *Informationsvermittlung* erbringen so gut wie alle informationsorientierten Informationsfunktionen einen mehr oder weniger großen informationellen Mehrwert in Form einer stärkeren fachlichen Informationsaufbereitung. Wenn diese Abteilungen nicht direkt der Unternehmensleitung unterstellt sind, sind sie verschiedenen Fachabteilungen zugeordnet.

²¹ Bei einem Vergleich von informations- und technologieorientierten Informationsfunktionen muss deren unterschiedliche Größe (Mitarbeiterzahl und Budgethöhe) berücksichtigt werden.

Die durchschnittliche Mitarbeiterzahl einer EDV-/Organisationsabteilung²² beträgt in den Unternehmen der KFZ-Industrie 45. Ihre Ermittlung war für die Banken und Versicherungen nur näherungsweise möglich. Sie liegt in diesen beiden Branchen aber deutlich über 100 Mitarbeiter. Auch bei den informationsorientierten Informationsfunktionen war die Berechnung dieser Kennzahl schwierig, nicht zuletzt deshalb, weil nicht sichergestellt werden kann, dass in den untersuchten Unternehmen alle Stellen identifiziert wurden. In sechs Unternehmen konnte nicht einmal *eine* informationsorientierte Informationsfunktion ermittelt werden. In vier davon handelt es sich um ausländische Unternehmen, bei denen die entsprechenden Stellen im Konzernsitz angesiedelt sind. Wenn man vom Erhebungsproblem absieht, besteht die durchschnittliche informationsorientierte Informationsfunktion aus 11 Mitarbeitern.

3.2 Strategische Informations(technologie)planung

A) Technologieorientierte Informationsfunktionen

Tabelle 2 gibt einen Überblick über die strategische IT-Planung. Wie zu erkennen ist, werden die Kriterien von einem Großteil der untersuchten Unternehmen erfüllt. Dies trifft besonders auf das Vorhandensein von *Verzeichnissen* zu, in denen die *Informationsressourcen* verwaltet werden. Bei den technologieorientierten Informationsfunktionen handelt es sich dabei primär um Hardware, Software, Lizenzen, Daten und Endbenutzer. Derartige Verzeichnisse sind in allen Unternehmen und bis auf einer in allen Informationsfunktionen vorhanden. Es werden von allen Informationsfunktionen, die Angaben dazu gemacht haben, Werkzeuge dafür verwendet. Es handelt sich oft um Data-Dictionaries, zum Teil werden selbst erstellte Informationssysteme, die zum Beispiel mit Lotus Notes entwickelt wurden, eingesetzt.

In immerhin einem Viertel der EDV-/Organisationsabteilungen gibt es keinen schriftlichen *Informationstechnologieplan*. Dies wurde mit „(Zwang zur) Orientierung an den gerade aktuellen Bedürfnissen“ und „Zeitmangel“ begründet. Im Falle einer IT-Planung wird diese ernsthaft wahrgenommen: In 14 Informationsfunktionen gibt es einen operativen, in 13 einen strategischen Informationsplan. In allen 15 Informationsfunktionen (für die diese Frage relevant war) basiert der IT-Plan auf den Unternehmenszielen und der Unternehmensstrategie. Geringer ist das Ausmaß, in dem die Leiter direkt an der Unternehmensplanung mitwirken: bei 11 der 20 Informationsfunktionen bzw. 9 der 16 Unternehmen. Bei der Studie von O'Brien/Morgan war die Zustim-

²² Eine ausgelagerte EDV-Abteilung wurde dabei wie eine interne behandelt.

mung bei einer ähnlich gestellten Frage etwas höher (68 %).²³ In den Banken waren obige Indikatoren deutlich stärker ausgeprägt.

	Banken (n=6 IF) ²⁴	Versiche- rungen (n=9 IF)	KFZ- Industrie (n=5 IF)	gesamt (n=20 IF)
Verzeichnis der Informationsressourcen (<i>Anzahl IF</i>)	6 (=100 %)	8 (=89 %)	5 (=100 %)	19 (=95 %)
Verzeichnis - Werkzeug (<i>Anzahl IF</i>)	5 (=100 %) (n=5) ²⁵	8 (=100 %) (n=8)	5 (=100 %)	18 (=100 %) (n=18)
schriftlicher Informationstechnologieplan (<i>Anzahl IF</i>)	6 (=100 %)	6 (=67 %)	3 (=60 %)	15 (=75 %)
operativer Informationstechnologieplan	6 (=100 %)	5 (=100 %) (n=6)	3 (=100 %) (n=3)	14 (=93 %) (n=15)
strategischer Informationstechnologieplan	4 (=67 %)	6 (=100 %) (n=6)	3 (=100 %) (n=3)	13 (=87 %) (n=15)
Abstimmung des IT-Plans mit Unternehmenszielen / -strategie	6 (=100 %)	6 (=100 %) (n=6)	3 (=100 %) (n=3)	15 (=100 %) (n=15)
direkte Mitarbeit bei Unternehmensplanung (<i>Anzahl IF</i>)	5 (=83 %)	4 (=44 %)	2 (=40 %)	11 (=55 %)
wettbewerbsorientierter Einsatz von IT (<i>Anzahl IF</i>)	6 (=100 %)	9 (=100 %)	5 (=100 %)	20 (=100 %)
Produktdifferenzierung	3 (=50 %)	7 (=78 %)	0 (=0 %)	10 (=50 %)
Kundendienst/Service	5 (=83 %)	8 (=89 %)	3 (=60 %)	16 (=80 %)
neue Produkte	6 (=100 %)	8 (=89 %)	1 (=20 %)	15 (=75 %)
Lieferantenbeziehungen	2 (=33 %)	0 (=0%)	4 (=80 %)	6 (=30 %)
Kostensenkung	5 (=83 %)	9 (=100 %)	5 (=100 %)	19 (=95 %)
Marktnischen	6 (=100 %)	3 (=33 %)	0 (=0 %)	9 (=45 %)
Zusammenarbeit mit anderen Unternehmen	3 (=50 %)	3 (=33 %)	4 (=80 %)	10 (=50 %)

Tabelle 2: Strategische Informationstechnologieplanung - Ergebnisse technologieorientierte Informationsfunktionen (IF)

In allen Unternehmen wird versucht, durch den Einsatz von *Informationstechnologie Wettbewerbsvorteile* zu erreichen. Bezüglich des Schwerpunkts gibt es aber erhebliche Abweichungen vor allem zwischen Dienstleistungs- und Industrieunternehmen. In der Automobilindustrie gibt es traditionell eine enge Kooperation mit Lieferanten und anderen Unternehmen. Es ist daher

²³ Vgl. O'Brien/Morgen 1991.

²⁴ Diese Angabe (n=6 Informationsfunktionen) bezieht sich auf die Anzahl der „normalerweise“ vorliegenden Aussageeinheiten.

²⁵ Falls die Anzahl der Aussageeinheiten bei einzelnen Fragen abweicht (weil zum Beispiel die jeweilige Frage nicht für alle Interviewpartner relevant war), wurde dies explizit aufgezeigt.

nicht überraschend, dass in diesen beiden Punkten die Zustimmung von KFZ-Unternehmen hoch war. Banken und Versicherungen lassen sich dadurch charakterisieren, dass die Informationsintensität des Produkts hoch ist. Dadurch ist die Hebelwirkung des Informationsmanagements²⁶ hier höher als in der Automobilindustrie. Das ist eine mögliche Erklärung dafür, dass bei den Dienstleistungsunternehmen dem EDV-Einsatz bei der „Produktdifferenzierung“, der „Besetzung von Marktnischen“ und der „Einführung von neuen Produkten“ (zum Beispiel „Internet-Banking“) eine größere strategische Bedeutung zukommt. Insgesamt (über alle Branchen) am meisten Nennungen erhielten „Kostensenkung“ (19 bzw. 95 %) und „Kundendienst/Service“ (16 bzw. 80 %).

B) Informationsorientierte Informationsfunktionen

Aus der Gegenüberstellung von Tabelle 2 und 3 geht hervor, dass die Merkmalsausprägungen bei den informationsorientierten Informationsfunktionen geringer sind. Ein *Verzeichnis*, in dem die *Informationsressourcen* verwaltet werden, gibt es in 17 der 22 Informationsfunktionen. Aufgrund der Heterogenität der informationsorientierten Informationsfunktionen sind die Inhalte teilweise recht unterschiedlich. Beispiele dafür sind Verweise auf Experten, Datenbanken, Kundendaten, Handbücher, Zeitschriften, Bücher, Akte, selbst erstellte Dokumente oder Inhaltsverzeichnisse zu in digitaler Form vorliegenden Daten. 13 Informationsfunktionen setzen ein Werkzeug zur Verwaltung der Informationsressourcen ein. Öfters handelt es sich um mit Lotus Notes oder MS Access selbst erstellte Informationssysteme. Mitunter werden die Informationsressourcen mit einem Tabellenkalkulationsprogramm verwaltet.

Einen schriftlich ausformulierten *Informationsplan* gibt es in genau der Hälfte der Informationsfunktionen. In jedem dieser Fälle wird ein operativer Informationsplan erstellt, in weniger als der Hälfte davon (fünf) ein strategischer Informationsplan. Neun Interviewpartner gaben an, dass die Informationsplanung auf den Unternehmenszielen basiert; ebenso viele, dass sie bei der Unternehmensplanung direkt mitarbeiten. Im Falle einer fehlenden Informationsplanung wurden folgende Gründe genannt: „Stelle erst im Aufbau begriffen“, „zu wenig Zeit und Ressourcen“, „fehlende Struktur im Unternehmen“, „Schwerpunkt liegt bei Informationsanalyse und –aufbereitung“, „alles im Kopf“ sowie „Stelle von der Auflösung bedroht“.

Fast alle Befragten (20 bzw. 91 %) stimmten zu, dass in ihrem Unternehmen aktiv versucht wird, durch den Einsatz von *Information Wettbewerbsvorteile* zu erreichen. Am öftesten (16-mal) wurde „Kundendienst/Service“ genannt.

²⁶ Vgl. zum Beispiel Picot/Franck 1992, S. 888.

Einer Kostensenkung wird geringere Bedeutung beigemessen. Die hier vorliegende Untersuchung bestätigt somit die Ergebnisse der Studie von Olaisen²⁷, wonach es Unterschiede zwischen dem strategischen Einsatz von Information und jenem von Informationstechnologie gibt.

	Banken (n=13 IF)	Versicherungen (n=5 IF)	KFZ-Industrie (n=4 IF)	gesamt (n=22 IF)
Verzeichnis der Informationsressourcen (<i>Anzahl IF</i>)	11 (=92 %) (n=12)	2 (=40 %)	4 (=100 %)	17 (=81 %)
Verzeichnis - Werkzeug (<i>Anzahl IF</i>)	10 (=100 %) (n=10)	0 (=0 %) (n=2)	4 (=100 %)	14 (=88 %) (n=16)
Schriftlicher Informationsplan (<i>Anzahl IF</i>)	6 (=46 %)	1 (=20 %)	4 (=100 %)	11 (=50 %)
operativer Informationsplan	6 (=100 %)	1 (n=1)	4 (=100 %)	11 (=100 %) (n=11)
strategischer Informationsplan	3 (=50 %) (n=6)	0 (n=1)	2 (=50 %)	5 (=45 %) (n=11)
Abstimmung des Informationsplans mit Unternehmenszielen/-strategie	4 (=67 %) (n=6)	1 (n=1)	4 (=100 %)	9 (=82 %) (n=11)
Direkte Mitarbeit bei Unternehmensplanung (<i>Anzahl IF</i>)	4 (=31 %)	2 (=40 %)	3 (=75 %)	9 (=41 %)
Wettbewerbsorientierter Einsatz von Information (<i>Anzahl IF</i>)	12 (=92 %)	5 (=100 %)	3 (=75 %)	20 (=91 %)
Produktdifferenzierung	8 (=67 %) (n=12)	2 (=40 %)	1 (=33 %) (n=3)	11 (=55 %) (n=20)
Kundendienst/Service	9 (=75 %) (n=12)	4 (=80 %)	3 (=100 %) (n=3)	16 (=80 %) (n=20)
neue Produkte	7 (=58 %) (n=12)	4 (=80 %)	3 (=100 %) (n=3)	14 (=70 %) (n=20)
Lieferantenbeziehungen	1 (=8 %) (n=12)	0 (=0%)	1 (=33 %) (n=3)	2 (=10 %) (n=20)
Kostensenkung	6 (=50 %) (n=12)	1 (=20 %)	2 (=67 %) (n=3)	9 (=45 %) (n=20)
Marktnischen	9 (=75 %) (n=12)	1 (=20 %)	2 (=67 %) (n=3)	12 (=60 %) (n=20)
Zusammenarbeit mit anderen Unternehmen	2 (=17 %) (n=12)	1 (=20 %)	2 (=67 %) (n=3)	5 (=25 %) (n=20)

Tabelle 3: Strategische Informationsplanung – Ergebnisse informationsorientierte Informationsfunktionen (IF)

²⁷ Vgl. Olaisen 1990.

3.3 Informationsrichtlinien

A) Technologieorientierte Informationsfunktionen

	Banken (n=6 IF)	Versicherungen (n=9 IF)	KFZ-Industrie (n=5 IF)	gesamt (n=20 IF)
Informationsrichtlinien (Anzahl IF)	schriftlich: 5 „im Kopf“: 1	schriftlich: 9	schriftlich: 5	schriftlich: 19 „im Kopf“: 1
Bekanntheitsgrad (Median)	größtenteils	überall	größtenteils	überall
Bindung daran (Median)	größtenteils	größtenteils	größtenteils	größtenteils

Tabelle 4: Informationsrichtlinien – Ergebnisse technologieorientierte Informationsfunktionen (IF)

Bis auf eine gibt es in allen anderen technologieorientierten Informationsfunktionen schriftliche Informationsrichtlinien. Dies kann als ein Indikator für die Professionalität dieser Abteilungen angesehen werden. Die Informationsrichtlinien können sehr umfangreich sein, was auch durch die große Breite der Antworten zum Ausdruck kam. Unter anderem wurden folgende Inhalte genannt: Datenschutz und –sicherheit, Beschaffungsrichtlinien, Abschreibungsdauer für Hardware, PC-Einsatz, Internet-Verwendung, Umgang mit Daten oder Anwendungsentwicklung. Die Informationsrichtlinien sind in unterschiedlichen Dokumenten festgehalten. Es kann sich dabei zum Beispiel um Dienstverträge, Arbeitsanweisungen, Organisationshandbücher, ISO-Dokumente, Handbücher für die Softwareentwicklung, Werksnormen oder Bereichsrichtlinien der Informatik handeln. Oft sind die Richtlinien im Internet oder in Lotus Notes-Datenbanken veröffentlicht. In der Regel sind die Richtlinien größtenteils oder sogar beim gesamten dafür in Frage kommenden Personenkreis bekannt. Sie werden auch größtenteils eingehalten. Ein Befragter, in dessen Unternehmen die Bindung an Informationsrichtlinien geringer ist, wies auf das Problem hin, dass es nach einer erfolgten Fusion Probleme mit unterschiedlichen „Standards“ geben kann.

B) Informationsorientierte Informationsfunktionen

Von nur 59 % der informationsorientierten Informationsfunktionen werden schriftliche Informationsrichtlinien herausgegeben. Die Inhalte der Informationsrichtlinien unterscheiden sich klarerweise von jenen der technologieorientierten Informationsfunktionen. Sie betreffen unter anderem Bibliotheksbenutzung, kostengünstige Durchführung von Datenbankrecherchen, Geheimhaltung von Daten, Umgang mit vertraulichen Daten, Berechtigungen bezüglich Informationsverteilung und dazu zu verwendende Medien, Zeitpunkt ab wann neue Informationen jeden Tag im Intranet verfügbar sein müssen, Er-

scheinungsbild der Dokumente, Abstimmung von interner und externer Kommunikation oder Kriterien, die bei der Informationsproduktion zu beachten sind. Die Bedeutung von Informationsrichtlinien ist im Bankenwesen am Bankgeheimnis erkennbar. Nicht zuletzt deshalb herrscht in dieser Branche eine eher restriktive Informationskultur vor. Falls vorhanden, ist der Bekanntheitsgrad der Informationsrichtlinien sehr hoch (Median: überall bekannt), die Bindung daran naturgemäß etwas niedriger (Median: größtenteils eingehalten).

	<i>Banken</i> (n=13 IF)	Versicherungen (n=5 IF)	KFZ-Industrie (n=4IF)	gesamt (n=22 IF)
Informationsrichtlinien (<i>Anzahl IF</i>)	schriftlich: 6 „im Kopf“: 5 keine: 2	schriftlich: 3 keine: 2	schriftlich: 4	schriftlich: 13 „im Kopf“: 5 keine: 4
Bekanntheitsgrad (<i>Median</i>)	überall (n=8)	überall (n=3)	größtenteils/überall	überall
Bindung daran (<i>Median</i>)	größtenteils (n=8)	größtenteils (n=3)	größtenteils	größtenteils

Tabelle 5: Informationsrichtlinien – Ergebnisse informationsorientierte Informationsfunktionen (IF)

3.4 Informations(verarbeitungs)controlling

A) Technologieorientierte Informationsfunktionen

Die Informationstechnologie ist ein bedeutender Kostenfaktor. Dies kommt in der hier vorliegenden Studie dadurch zum Ausdruck, dass in einigen Unternehmen ein *EDV-Budget* von mehreren zig-Millionen Euro erhoben wurde. Ein exakter Branchenvergleich ist aber nicht möglich, da aufgrund von Ausreißern bei den Banken und Versicherungen die Anonymität einzelner Unternehmen nicht gewährleistet werden hätte können und nur die Hälfte der Unternehmen dieser beiden Branchen ihr EDV-Budget bekannt gab. Hingegen machten alle Unternehmen der KFZ-Industrie dazu eine Angabe. Demnach beläuft sich das durchschnittliche EDV-Budget hier auf zirka sieben Millionen Euro. Ein Mehrfaches davon betragen die EDV-Budgets in der Banken- und Versicherungsbranche, wobei die Banken am meisten Geld für Informationstechnologie ausgeben.

In der Literatur wird eine möglichst verursachungsgerechte *Verrechnung* der Kosten der Informationsverarbeitung gefordert.²⁸ Die untersuchten Informationsfunktionen versuchen, dem auch Rechnung zu tragen. Genau die Hälfte verwendet Verrechnungspreise zu Vollkosten, eine Informationsfunktion verrechnet ihre Leistungen zu Teilkosten. Eine Kostenstellenumlage gibt es nur

²⁸ Vgl. Horvath 1998; Spitta 1998, S. 429.

in zwei Unternehmen. In einem Viertel der Informationsfunktionen kommen Mischsysteme aus obigen Kostenverrechnungsverfahren zum Einsatz. In zwei Unternehmen werden die anfallenden Kosten als unternehmensfix betrachtet.

	<i>Banken</i> (n=6 IF)	Versicherungen (n=9 IF)	KFZ-Industrie (n=5 IF)	gesamt (n=20 IF)
Cost/Profit Center, Sonstiges (Anzahl IF)	Cost Center: 4 Sonstiges: 2	Cost Center: 9	Cost Center: 3 Profit Center: 2	Cost Center: 16 Profit Center: 2 Sonstiges: 2
Kennzahlen (Anzahl IF)	6 (=100 %)	8 (=89 %)	4 (=80 %)	18 (=90 %)
Kostenplanung und -steuerung (Anzahl IF)	6 (=100 %)	8 (=80 %)	3 (=60 %)	17 (=85 %)
Kostenverrechnung (Anzahl IF)				
Verrechnungspreise	4 (=67 %)	3 (=33 %)	4 (=80 %)	11 (=55 %)
Zu Vollkosten	3 (=50 %)	3 (=33 %)	4 (=80 %)	10 (=50 %)
Zu Teilkosten	1 (=17 %)	0	0	1 (=5 %)
Kostenstellenumlage	0	2 (=22 %)	0	2 (=10 %)
Unternehmensfixe Kosten	0	1 (=11 %)	1 (=20 %)	2 (=10 %)
Mischsysteme	2 (=33 %)	3 (=33 %)	0	5 (=25 %)

Tabelle 6: Informationsverarbeitungs-Controlling: Ergebnisse technologieorientierte Informationsfunktionen (IF)

16 EDV-Abteilungen sind als *Cost Center* und nur zwei als *Profit Center* eingerichtet. Es handelt sich dabei um zwei ausgelagerte EDV-Abteilungen. 17 Interviewpartner (85 %) gaben an, dass sie die Kosten ihrer Informationsfunktion *nach Kostenarten planen und steuern*. Obwohl die einzelnen Informationsfunktionen teilweise von unterschiedlichen Kostensystematiken ausgehen, lässt sich sagen, dass die Personalkosten normalerweise den höchsten Kostenfaktor bilden. In den meisten technologieorientierten Informationsfunktionen machen Kosten für Hardware und Anwendungssoftware ebenfalls einen großen Anteil an den IT-Gesamtkosten aus.

Kennzahlen werden von fast allen Informationsfunktionen (90 %) ermittelt. Im Bereich des Infrastruktur-Controlling sind dies unter anderem Bestandskennzahlen wie Anzahl der Computer, Anzahl der Applikationen, Anzahl der Endbenutzerwerkzeuge, Datenbankanzahl, Terminalanzahl oder EDV-Aufwand je Mitarbeiter. Die Prozesskennzahlen beziehen sich hauptsächlich auf herkömmliche Rechenzentrumsstatistiken: durchgeführte Transaktionen, Datenbankzugriffe, Antwortzeiten, Durchlaufzeit von EDV-Aufträgen, Ausfallsstatistiken oder Rechnerauslastung. Stärker am Endbenutzer oder Unternehmen orientierte Kennzahlen, zum Beispiel Mitarbeiterzufriedenheit oder Beschleunigung der Durchlaufzeit von Aufträgen (durch den EDV-Einsatz),

werden seltener ermittelt. Ein EDV-Leiter meinte in diesem Zusammenhang, dass es primär diese Kennzahlen sind, die für ihn von Interesse sind.

B) Informationsorientierte Informationsfunktionen

	Banken (n=13 IF)	Versicherungen (n=5 IF)	KFZ-Industrie (n=4 IF)	gesamt (n=22 IF)
Cost/Profit Center, Sonstiges (Anzahl IF)	Cost Center: 9 Profit Center: 1 Sonstiges: 3	Cost Center: 4 (n=4)	Cost Center: 3 Sonstiges: 1	Cost Center: 16 Profit Center: 1 Sonstiges: 4 (n= 21)
Durchschnittliches Budget (Mittelwert in 1000 €)	1.090 € (n=10)	174 € (n=1)	1.744 €	1.163 € (n=15)
Kennzahlen (Anzahl IF)	7 (=54 %)	2 (=40 %)	4 (= 100 %)	13 (= 59 %)
Kostenplanung und –steuerung (Anzahl IF)	7 (=58 %) (n=12)	2 (=40 %)	3 (= 75 %)	12 (=57 %) (n=21)
Kostenverrechnung (Anzahl IF)				
Verrechnungspreise	2 (=15 %)	0	2 (=50 %)	4 (=18 %)
zu Vollkosten	2 (= 15 %)	0	1 (=25 %)	3 (= 14 %)
zu Teilkosten	1 (= 8 %)	0	1 (= 25 %)	2 (= 9 %)
Kostenstellenumlage	3 (= 23 %)	0	2 (=50 %)	5 (= 23 %)
unternehmensfixe Kosten	7 (=54 %)	5 (=100 %)	0	12 (=50 %)
Mischsysteme	1 (= 8 %)	0	0	1 (= 5 %)

Tabelle 7: Informationscontrolling – Ergebnisse informationsorientierte Informationsfunktionen (IF)

Bei den informationsorientierten Informationsfunktionen macht das durchschnittlich *Budget* mit etwas mehr als einer Million Euro nur einen Bruchteil des Budgets der EDV-Abteilungen aus. Der Großteil der informationsorientierten Informationsfunktionen (16) wird ebenfalls als *Cost Center* geführt. Bei einer Abteilung handelt es sich um ein *Profit Center*. Diese bietet spezialisierte Informationsdienste an, die über die traditionelle Informationsvermittlung weit hinausgehen. Ein weiterer Interviewpartner hatte ebenfalls die Präferenz in Richtung Profit Center. Demnach sollten die für das Unternehmen erzielten Einnahmen der Informationsfunktion gutgeschrieben werden. Auf die damit verbundenen Quantifizierungsprobleme ging der Leiter dieser Abteilung aber nicht ein.

In zirka der Hälfte der informationsorientierten Informationsfunktionen wird eine *Kostenplanung und –steuerung* durchgeführt. Bis auf einen Interviewpartner identifizierten alle anderen die Personalkosten als den höchsten Kostenfaktor. Im Gegensatz zu den EDV-Abteilungen wurden die Informationsbeschaffungskosten von den meisten Abteilungen an die zweite Stelle gereiht.

Bei der *Weiterverrechnung* der anfallenden *Kosten* ergeben sich wieder große Unterschiede zu den technologieorientierten Informationsfunktionen: In etwas mehr als der Hälfte der informationsorientierten Informationsfunktionen (12) werden die anfallenden Kosten als unternehmensfix betrachtet.

In immerhin 18 Informationsfunktionen (90 %) werden *Kennzahlen* ermittelt, wenn auch die Detailliertheit der Kennzahlensysteme von Abteilung zu Abteilung stark variieren dürfte. Beispiele für verwendete Bestandskennzahlen sind Anzahl der Zeitschriftenabonnements, Literaturbestand, Anzahl der lizenzierten Datenbanken oder Bestand an Datensätzen bei elektronischen Datenbeständen. Übliche Prozesskennzahlen sind Benutzeranzahl, Benutzerstruktur nach Abteilungen, durchgeführte Anfragen/Aufträge (Anzahl, Zeitdauer, inhaltliche Verteilung nach Themen, Form der Beantwortung, durchschnittliche Kosten, ...), durchgeführte Datenbankrecherchen oder Anzahl der erstellten Dokumente.

3.5 Methoden

Wie aus Abbildung 2 ersichtlich ist, ist der Methodeneinsatz in den informationsorientierten Informationsfunktionen deutlich geringer als in den technologieorientierten. Ein Grund dürfte darin liegen, dass vom Methodeneinsatz großteils besser strukturierbare Bereiche betroffen sind. Im Bereich der strategischen Planung werden Methoden generell selten verwendet.

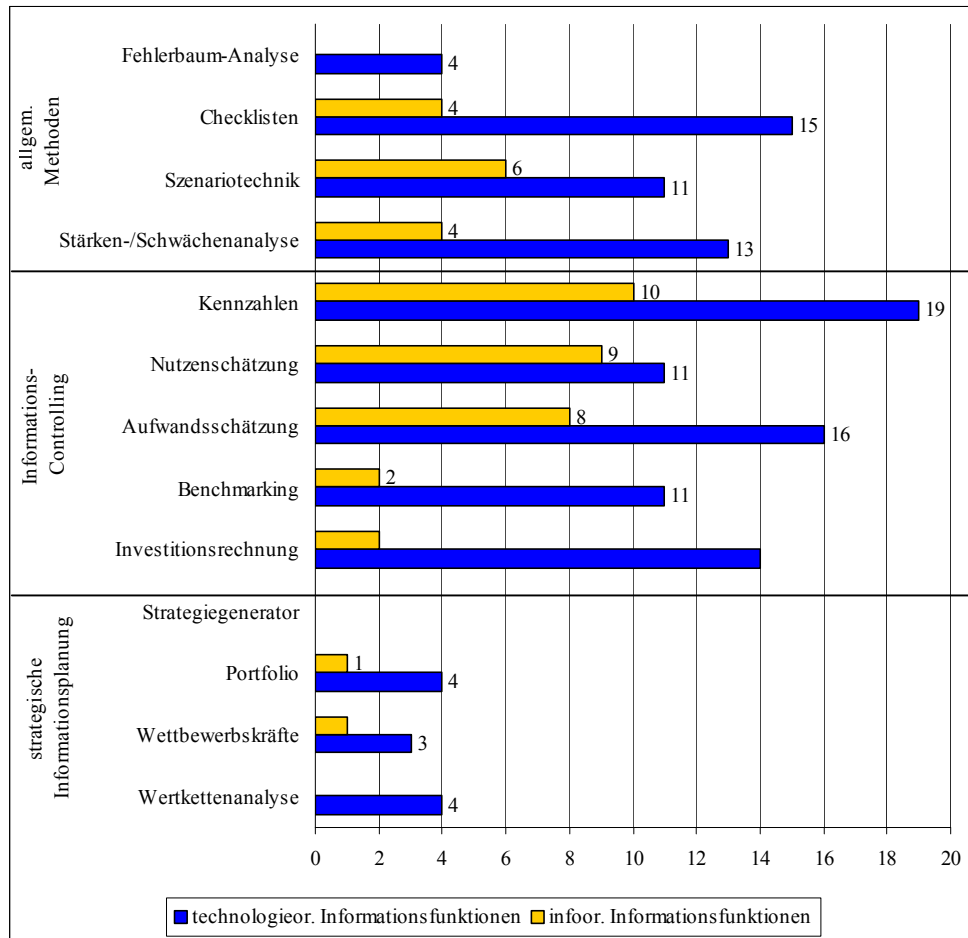


Abbildung 2: Methodeneinsatz – Gegenüberstellung von informations- und technologieorientierten Informationsfunktionen

3.6 Management

Unternehmensweite *IT-Standards* sind in allen Unternehmen vorhanden. Sie werden auch großteils von dem sie betreffenden Personenkreis eingehalten. Ebenfalls in allen Unternehmen gibt es eine Informationsarchitektur, wenn auch die einzelnen Komponenten in einem unterschiedlichen Maße realisiert sind. Am stärksten ausgebaut ist die Datenarchitektur (Median: großteils). Beim Realisierungsgrad ergeben sich kaum branchenmäßige Unterschiede. Lediglich die Prozessarchitektur ist in der Automobilindustrie vor allem im Vergleich zu den Banken weiter fortgeschritten. Da in der KFZ-Industrie Produktionsprozesse und –abläufe naturgemäß eine zentrale Rolle spielen, erscheint dieses Ergebnis plausibel. Neben einer Datenarchitektur ist die strukturenorganisatorische Verankerung nicht nur ein wichtiger Indikator des Datenmanagements, sondern auch des Informationsmanagements. Dem wird von den meisten Unternehmen auch entsprochen. Nur in einem Unternehmen der

KFZ-Industrie und in zwei Versicherungen gibt es weder eine Stelle noch eine (Sub)Abteilung für das *Datenmanagement*.

	Banken (n=4 Unternehmen)	Versicherungen (n=7 Unternehmen)	KFZ-Industrie (n=5 Unternehmen)	gesamt (n=16 Unternehmen)
IT-Standards (Anzahl Unternehmen)	4 (=100 %)	7 (=100 %)	5 (=100 %)	16 (=100 %)
Standards eingehalten (0 – 3 bzw. nein – ja) (Median)	eher ja (2)	ja (3)	ja (3)	ja (3)
Informationsarchitektur (Anzahl Unternehmen)	4 (=100 %)	7 (=100 %)	4 (=100 %) (n=4)	15 (=100 %)
Datenarchitektur (0 – 3 bzw. keine, teilweise, großteils, vollständig) (Median)	großteils (2)	großteils (2)	großteils (2)	großteils (2)
Funktionsarchitektur (Median)	großteils (2)	teilweise (1)	teilweise (1)	teilweise (1)
Prozessarchitektur (Median)	teilweise (1)	großteils (2)	großteils (2)	teilweise (1)
Stelle Datenmanagement (Anzahl Unternehmen)	4 (=100 %)	5 (=71 %)	4 (=80 %)	13 (=81 %)
Anteil Standardsoftware (Mittelwert)	29 % (n=3)	34 %	42 %	36 %
IT-Monitoring und Ausnutzen neuer Möglichkeiten (Anzahl Unternehmen)	3 (=75 %)	5 (=71 %)	4 (=80 %)	12 (=75 %)

Tabelle 8: IT-Management - Ergebnisse

Eine wichtige Aufgabe des IT-Managements besteht darin, die Entwicklung der Informationstechnologie laufend zu verfolgen und zu versuchen, die sich daraus ergebenden Chancen für das eigene Unternehmen zu nutzen. Dreiviertel der Unternehmen führen ein derartiges *Monitoring des IT-Marktes* durch. Ein EDV-Leiter gab sogar an, dass es sich dabei um eine seiner Hauptaufgaben handelt. Intensität und Form sind von Fall zu Fall verschieden und hängen unter anderem von Größe und Internationalität des Unternehmens ab.

4 Schlussfolgerungen

Aus obigen Ausführungen lassen sich folgende Schlussfolgerungen ziehen: In genau der Hälfte der Unternehmen ist die EDV-Abteilung ausgelagert. Ein *Outsourcing* stellt bei informationsorientierten Informationsfunktionen die Ausnahme dar. Wenn diese Stellen/Abteilungen nicht direkt der Unternehmensleitung unterstellt sind, sind sie verschiedenen Fachbereichen zugeordnet, was ihre Lokalisierung oft schwierig macht. Neben der reinen Informationsvermittlung sind diese Stellen in einem stärkeren Maße auch mit inhaltlicher Informationsanalyse und –aufbereitung betraut. Teilweise sind die Grenzen zu den Fachabteilungen fließend.

Der Bereich der *strategischen Informations(technologie)planung* wird in den untersuchten Unternehmen ernst genommen. In den meisten Unternehmen gibt es einen operativen und einen strategischen Informationstechnologieplan. In ungefähr der Hälfte der Unternehmen arbeitet der Leiter der EDV-/Organisationsabteilung bei der Unternehmensplanung mit. Hingegen erstellt nur zirka die Hälfte der informationsorientierten Informationsfunktionen einen operativen und ein Viertel einen strategischen Informationsplan. In fast allen Unternehmen wird versucht, durch den Einsatz von Information und Informationstechnologie Wettbewerbsvorteile zu erreichen. Dabei gibt es deutliche Unterschiede zwischen informations- und technologieorientierten Informationsfunktionen. Während bei EDV-/Organisationsabteilungen neben einer Verbesserung von Kundendienst bzw. Servicequalität vor allem Kosteneinsparungen dominieren, liegt der Fokus bei informationsorientierten Informationsfunktionen stärker auf der Verbesserung von Kundendienst/Service, der Einführung von neuen Produkten sowie dem Verfolgen einer Marktnischen- bzw. Differenzierungsstrategie.

Bis auf eine geben alle anderen technologieorientierten Informationsfunktionen schriftliche *Informationsrichtlinien* heraus. Dies trifft jedoch nur auf etwas mehr als die Hälfte der informationsorientierten Informationsfunktionen zu. Bei vorhandenen Richtlinien sind Bekanntheitsgrad und Bindung daran recht hoch.

Auch im Bereich des *Informations(verarbeitungs)controlling* gibt es deutliche Unterschiede zwischen informations- und technologieorientierten Informationsfunktionen. Eine Kostenplanung und –steuerung erfolgt bei fast allen technologieorientierten und bei mehr als der Hälfte der informationsorientierten Informationsfunktionen. Große Unterschiede gibt es auch bei der Leistungsverrechnung. Während bei den EDV-Abteilungen das Bemühen nach einer möglichst verursachungsgerechten Kostenverrechnung (Verrechnungspreise, Mischsysteme) besteht, werden die anfallenden Kosten bei den meisten informationsorientierten Informationsfunktionen als unternehmensfix betrachtet. Der mit der Kostenverrechnung verbundene Aufwand ist natürlich bei einem hohen Informationsbudget eher gerechtfertigt. Während dieses bei informationsorientierten Informationsfunktionen bei über einer Million Euro liegt, beläuft es sich bei den EDV-Abteilungen im Durchschnitt auf mehrere zig-millionen Euro. Der Großteil der Informationsfunktionen wird als Cost Center geführt. Nur zwei EDV-Abteilungen wurden als Profit Center eingerichtet.

Das Ausmaß der *methodischen Unterstützung* ist bei den technologieorientierten Informationsfunktionen wesentlich höher. Bei der strategischen Informationsplanung ist der Methodeneinsatz aber auch von Seiten der EDV-/Organisationsabteilungen gering. Unternehmensweite IT-Standards sind in allen Unternehmen vorhanden und werden großteils eingehalten. Eine Informationsarchitektur gibt es ebenfalls in allen Unternehmen, wobei aber die einzelnen Komponenten in einem unterschiedlichen Ausmaß realisiert sind. Am weitesten ausgebaut ist die Datenarchitektur. Zu ihrer Betreuung ist in den meisten Unternehmen eine Stelle/(Sub)Abteilung Datenmanagement eingerichtet. In vielen Unternehmen erfolgt eine laufende Beobachtung des IT-Marktes hinsichtlich technologischer Neuerungen. Form und Intensität variieren aber von Unternehmen zu Unternehmen.

Wie obige Ausführungen darlegten, sind die einzelnen Bereiche des Informationsmanagements bei den technologieorientierten Informationsfunktionen deutlich stärker ausgeprägt als bei den informationsorientierten. Die Hauptursache dürfte in der unterschiedlichen Größe liegen. Zum Beispiel beträgt das Budget der informationsorientierten Informationsfunktionen im besten Fall ein Zehntel von jenem der EDV-/Organisationsabteilungen. Wenn man zusätzlich die schwache organisatorische Positionierung der informationsorientierten Informationsfunktionen berücksichtigt, die sich unter anderem daran äußert, dass diese in den einzelnen Unternehmen nur schwer identifizierbar sind, zeigt das die starke Betonung der technologieorientierten Informationsfunktionen im Rahmen des Informationsmanagements auf.²⁹ Der Hauptgrund liegt in der großen Bedeutung des IT-Einsatzes in der Wirtschaft. Die EDV stellt die Infrastruktur bereit, über die die Informationsvermittlung läuft. Diese wird aber mittlerweile zu einem großen Teil von den Nutzern in den Fachabteilungen selbsttätig wahrgenommen.

6 Literaturverzeichnis

- Bergeron P. (1995): An Examination of the Perceptions and Practices of Information Resources Management in Large Organizations from the Canadian Private Sector. Dissertation. Syracuse University, Syracuse.
- Bergeron P. (1996): Information Resources Management. In: Annual Review of Information Science and Technology, 31(1996), S. 263 - 300.
- Biethahn J., Mucksch H., Ruf W. (1996): Ganzheitliches Informationsmanagement. 4. Auflage. Oldenbourg, München.

²⁹ Zu einer ähnlichen Schlussfolgerung kommt auch Bergeron (vgl. Bergeron 1995, S. 456 f.).

- Brancheau J. C., Wetherbe J. C. (1987): Key Issues in Information Systems Management. In: MIS Quarterly, (1987) März, S. 23 - 45.
- Heinrich L. J. (1999): Informationsmanagement. 6. Auflage. Oldenbourg, München.
- Gemünden G., Schmitt M. (1991): Datenmanagement in deutschen Großunternehmen – Theoretischer Ansatz und empirische Untersuchung. In: Information Management, (1991) 4, S. 22 - 34.
- Horton F. W. (1985): Information Resources Management: Harnessing Information Assets for Productivity Gains in the Office, Factory and Laboratory. Prentice Hall, Englewood Cliffs.
- Horvath P. (1998): Controlling. 7. Auflage. Vahlen, München.
- Lehner F. (1993): Informatik-Strategien – Entwicklung, Einsatz und Erfahrungen. Hanser, München.
- Lewis D., Martin W. (1989): Information Management – State of the Art in the United Kingdom. In: ASLIB Proceedings, 41(1989)7-8, S. 225 - 250.
- Lewis B. R., Snyder C. A., Rainer R. K. (1995): An Empirical Assessment of the Information Resource Management Construct. In: Journal of Management Information Systems, 12 (1995) 1, S. 199 - 223.
- Lytle R. H. (1986): Information Resource-Management. In: Annual Review of Information Science and Technology, 21(1986), S. 309 - 336.
- Lytle R. H. (1988): Information Resource Management: A Five-Year Perspective. In: Information Management Review, 3(1988)3, S. 9 - 16.
- Marchand D. A., Horton F. W. (1986): Infotrends. Profiting from Your Information Resources. Wiley, New York.
- Niedereicholz J., Wentzel C. (1985): Voraussetzungen und organisatorische Wirkungen des Informationsmanagements. In: Angewandte Informatik, (1985)7, S. 284 - 290.
- Niederman F., Brancheau J. C., Wetherbe J. C. (1991): Information Systems Management Issues for the 1990's. In: MIS Quarterly, 15(1991)4, S. 475 - 500.
- Olaisen J. (1990): Information Versus Information Technology as a Strategic Resource: Areas of Application of Information and Information Technology in Norwegian Banks and Insurance Companies. In: International Journal of Information Management, 10(1990), S. 192 - 214.
- Orna E. (1990): Practical Information Policies. Quartus, London.
- O'Brien J. A., Morgan J. N. (1991): A Multidimensional Model of Information Resource Management. In: Information Resources Management Journal, 4(1991) 2, S. 2 - 11.
- Picot A., Franck E. (1992): Informationsmanagement. In: Frese E. (Hrsg.): Handwörterbuch der Organisation. 3. Auflage. Poeschel, Stuttgart, S. 886 - 900.
- Schlögl C. (2001): Bestandsaufnahme Informationsmanagement: Eine szientometrische, qualitative und empirische Analyse. Gabler, Wiesbaden.
- Schlögl C. (2002): Informationsvermittlung in österreichischen Großunternehmen. In: Password, (2000)4. S. 19 - 26.
- Schlögl C., Voglmayr B. (1999): Welche Informationsmanager braucht die Wirtschaft? In: Nachrichten für Dokumentation, 50(1999), S. 211 - 216.
- Spitta T. (1998): IV-Controlling in mittelständischen Industrieunternehmen – Ergebnisse einer empirischen Studie. In: Wirtschaftsinformatik, 40(1998)5, S. 424 - 433.

- Synnott W. R. (1987): *The Information Weapon: Winning Customers and Markets with Technology*. Wiley, New York.
- Taylor R. S. (1986): *Value-Added Processes in Information Systems*. Ablex, Norwood.
- Wigand R. T. (1988): Fünf Grundsätze für die erfolgreiche Einführung des Informations-Managements. In: *Information Management*, (1988)2, S. 24 - 30.



Information überall, jederzeit

Wolfgang Semar

Universität Konstanz
Fachbereich Informatik und Informationswissenschaft
Fach D 87
78457 Konstanz
Deutschland
wolfgang.semar@uni-konstanz.de

Zusammenfassung

Ubiquitär heißt „überall verbreitet“, und bezogen auf die entstehende Informationsgesellschaft wird dieses Attribut gerne herangezogen, um auf die Vision von PC und Internet in jedem Büro und in jedem Haushalt anzuspitzen. Informationstechnologien werden aber erst dann ubiquitär, wenn ihre Bindung an Standardrechner wie PCs überwunden wird und einer Einbettung in reale Umgebungen und Abläufe weicht. Dieser Beitrag zeigt auf, warum der PC ein Auslaufmodell ist und wie er durch einen neuen Entwicklungstrend, an dessen Anfang wir stehen, abgelöst wird. Im Mittelpunkt steht der Begriff des Information Appliance (IA).

1 Einleitung

Computer, die mit unseren alltäglichen Umgebungen verschmelzen und den PC heutigen Zuschnitts ablösen, wurden bereits Anfang der 90er Jahre von Mark Weiser [Weiser, Mark 1991] beschrieben. Technologien sind ein Mittel zum Zweck, die besten Technologien sind jene, die in den Hintergrund treten und mit den Abläufen, die sie unterstützen, eins werden. Weiser führt in seinem Grundsatzartikel als klassisches Beispiel Schrift an, die es im Gegensatz zu gesprochener Sprache erlaubt, Information festzuhalten und in neuartiger Form zu verteilen. Schrift ist ubiquitär, ist verwoben mit allen Lebensbereichen und tritt zu Gunsten der Anwendung - Speicherung und Kommunikation von Information - gänzlich in den Hintergrund. Moderne Informationstechnologien werden diesem Anspruch dort gerecht, wo sie für den Anwender unsichtbar eingebettet sind, beispielsweise in Haushaltsgeräten, Steuerungen



oder Leitsystemen. Der primäre Berührungspunkt mit modernen Informationstechnologien sind jedoch PCs und vergleichbare Standardrechner. Diese können diesem Anspruch nicht gerecht werden, wie Norman nachdrücklich aufzeigt [Norman, Donald A. 1998]. Aus diesem Grund ist der PC ein Auslaufmodell und wird durch einen neuen Entwicklungstrend, an dessen Anfang wir stehen, abgelöst. Die Technologien hierfür sind im Entstehen und in vielen Beispielen schon vorhanden. Im Mittelpunkt steht der Begriff des Information Appliance (IA), der anwendungsspezifische Geräte bezeichnet, die in der Lage sind, Information untereinander auszutauschen. Als Geräte, die für bestimmte Anwendungen entworfen wurden, fügen sich Information Appliances natürlicher in Abläufe ein als Standardrechner. Mit der Fähigkeit, Information untereinander austauschen zu können, leisten Information Appliances dabei im Verbund eine dezentrale informationstechnische Unterstützung. Auf dem Weg zu vernetzten Information Appliances entstehen zum einen innovative persönliche Technologien und zum anderen informationstechnisch erschlossene Umgebungen (Informationsumgebung). Neue persönliche Technologien sehen vor allem vor, über eine neue Qualität der Portabilität von Informationstechnologien ubiquitären Zugriff auf Informationsdienstleistungen zu erreichen, ubiquitär steht hier für IT, die überall mithin genommen werden kann. In Informationsumgebungen ist der Ansatz hingegen, Informationstechnologien in die räumliche Umgebung einzubetten, hier steht ubiquitär für IT, die überall vorhanden ist.

2 Vom PC zum IA

Heutzutage können wir kaum noch eine Arbeit erledigen, ohne dabei auf die Informationstechnologie zu verzichten. Aber erfüllt der PC, den wir dabei benutzen, wirklich unsere Bedürfnisse? Die Computerindustrie ist heute an einem Punkt angekommen, an dem der Technologie-"Wahn" im Vordergrund steht. Mega-, Giga-, TeraByte und -hertz, das sind die Schlagworte, mit denen um neue Kunden geworben wird. Neue Software braucht immer schnellere und speicherintensivere Computer, jedes Update verlangt nach mehr und mehr Technologie. Aber der zukünftige Nutzer wird eher der pragmatische, konservative Kunde sein, der abgewartet hat, bis sich eine Technologie etabliert, bevor er sie ohne viele Mühe und Aufhebens nutzt. Dieses Kundensegment, die „late adopters“, ist nicht so technologie-orientiert wie das Kundensegment der „early adopters“¹.

¹ Siehe hierzu die Ergebnisse der empirischen Untersuchung von [Buse, Stephan 2002].

PCs sind Vielzweckwerkzeuge, ausgelegt als zentraler Zugang zur Informationsunterstützung verschiedenster Anwendungen und Abläufe, und von daher nicht geeignet, mit einzelnen Abläufen zu verschmelzen und dabei als Technologie in den Hintergrund zu treten. Es handelt sich hier nicht um ein Problem der Ergonomie der Rechner und darauf installierter Software, sondern um ein Problem der Zentralisierung informationstechnischer Unterstützung. Standardrechner unterstützen immer vielfältigere Aufgaben, neben den üblichen Arbeitsplatz- und Heimanwendungen beispielsweise Telekommunikation, Überwachung, Umgebungssteuerung und Empfang von Broadcastmedien. Standardrechner integrieren damit Funktionen anwendungsspezifischer Geräte wie beispielsweise Telefon, Fax, Radio und Haussteuerung mit dem Vorteil gebündelter informationstechnischer Unterstützung, aber dem Nachteil eines aus realen Abläufen herausgelösten und komplexen Zugangs. Ein neuer Entwicklungstrend, an dessen Anfang wir stehen, ermöglicht eine Zurückverlagerung heute in Standardrechnern gebündelter Funktionen in die eigentlichen Anwendungen, ohne die informationstechnische Unterstützung aufzugeben.

Zunächst ist der PC nicht sehr gebrauchsfähig, er ist eigentlich ein universales Werkzeug. Nur wenn ein Computer mit der entsprechenden Software ausgestattet wird wie z. B. Textverarbeitung, Tabellenkalkulation oder Email und er zusätzlich die notwendige Peripherie wie z. B. Speicherplatten, Drucker und Modem erhält, wird sein Nutzen größer. Das ist das traditionelle Computermodell. Die Arbeit, die der Anwender zu erledigen hat, geht dabei zu Lasten einer erhöhten Komplexität und Unbequemlichkeit. Das Tool selbst dominiert die Aktivität. Momentan ist es das Ziel der Computerindustrie, eine Maschine zu bauen und zu verkaufen, die alles kann und die mit Hilfe der gleichen Software und Hardware allen Menschen dienlich sein soll. Dieses ist zwar ein sehr schönes Geschäftsmodell für die Unternehmen, aber nicht für den Konsumenten.

Momentan befindet sich der PC immer noch in seinen Anfangstagen, sehr komplex, schwer zu bedienen und sehr teuer. Der Computer muss jetzt seinen Wandel zum Alltagsobjekt vollziehen, hin zu einem Instrument mit einfacher Bedienung und geringem Preis, nur so kann er die Welt der late adopters erobern. Wenn der konservative, pragmatische Kunde in den Markt eintritt, also die late adopters zu kaufen beginnen, dann tritt der Massenmarkt zum Kauf an. Ab jetzt funktioniert der altbekannte Geschäftsprozess nicht mehr. Diese Kunden wollen Zuverlässigkeit, Ausfallsicherheit, Beständigkeit, Bedienerfreundlichkeit, Einfachheit, ein attraktives Erscheinen sowie Prestige und einen Markenartikel. Ihr Credo ist: „Schalte es an, nutze es und beschäftige dich nicht mehr damit“. Von nun an ist es wichtig, dass das Produkt schnell und

einfach zu bedienen und preiswert ist. Es müssen endanwender-orientierte Produkte entwickelt werden, der Schwerpunkt auf die Technologie eines Produktes ist jetzt unangebracht. In dieser Phase fällt ein Produkt durch, auch wenn es eine hervorragende Technologie besitzt, diese jedoch schwer zu bedienen, zu nutzen oder zu teuer ist. Die besten Technologien sind jene, die in den Hintergrund treten und mit den Abläufen, die sie unterstützen, eins werden, dies führt zur Entwicklung der Information Appliances.

Das primäre Ziel der Information Appliances ist, die Komplexität der heutigen PCs zu brechen. Ein PC, der viele Dinge können muss, muss zwangsläufig Kompromisse bei jeder einzelnen Aufgabe eingehen. Entwickelt man aber ein Gerät für eine einzige Funktion, so kann es diese bestmöglich ausführen. Unter Information Appliances versteht man eine noch recht junge Kategorie von digitalen Endgeräten, die einen kostengünstigen, leicht handhabbaren, endanwender-orientierten Zugriff auf die verschiedensten Tätigkeiten ermöglichen. Information Appliances werden entwickelt, um eine ganz spezifische Aktivität zu unterstützen, wie z. B. Musik, Fotografie oder Schreiben. Jede einzelne Anwendung hat also ihr spezielles Gerät, das nur die Aufgabe unterstützt, für die es geschaffen wurde, und zwar so, als sei es eine natürliche Erweiterung des Anwenders. Obwohl ein einzelnes Gerät selbst sehr gut funktioniert, entwickelt sich die Leistungsstärke der IAs daraus, dass sie in einem System von interkonnektierenden Komponenten arbeiten. So ergibt sich die Leistungsstärke von elektronischen Musikinstrumenten auch erst durch ihre „Verschaltung“ untereinander. Wir brauchen also nicht einzelne Produkte, sondern eine ganze Produktfamilie, die miteinander mühelos und reibungslos arbeiten und kommunizieren können, direkt von Gerät zu Gerät und über das Internet. Wir nutzen ein Gerät für eine Funktion und schicken das Ergebnis per Knopfdruck zu einem anderen Gerät.

Für die Entwicklung sogenannter persönlicher Information Appliances sind zwei Entwicklungen entscheidend. Zum einen die Miniaturisierung, durch die sie kann die Technologie überall mithingenommen und eingesetzt werden, ohne als Belastung (oder sogar Ballast) empfunden zu werden. Für diese Eigenschaft wird gelegentlich der Begriff der Ultra-Portabilität benutzt, um eine neue Qualität der Portabilität zu beschreiben, die über die Portabilität beispielsweise von Notebook-Computern hinausgeht. Die Miniaturisierung wird ergänzt durch die Spezialisierung, um vom portablen Standardrechner zu persönlichen Information Appliances überzugehen. Vor dem Hintergrund der ab Anfang 2003 erwarteten Inbetriebnahme der UMTS-Netze entwickelt die Endgeräteindustrie zunehmend portable Systeme, die nicht mehr rein sprachorientiert sind. Hierzu gehören neben Mobiltelefon mit erweiterten Daten-

speicherfunktionen derzeit sogenannte Personal Digital Assistants (PDAs) [Keupner, Frank 2002], die vom Funktionsumfang noch relativ nah an Standardrechnern sind. Diese Lösungen sind aufgrund ihrer Bildschirmgröße, den Eingabemöglichkeiten sowie der Stromversorgung im Vergleich zu konventionellen Endgeräten allerdings nur begrenzt nutzbar [Schumann, Matthias 2002]. Mit Hilfe des sog. Wireless Application Protocol (WAP) können ein Teil der Mobiltelefone über Gateways auf speziell aufbereitete Inhalte im Internet zugreifen, wodurch der Benutzer praktisch jederzeit und überall erreichbar ist und eine zeitnahe Informationsübermittlung ermöglicht wird. Zeitkritische Daten können somit deutlich schneller empfangen werden, als mit einem stationären PC [Buse, Stephan 2002].

Eine besondere Form persönlicher Information Appliances sind Wearable Computers, die, so die Vision, so selbstverständlich wie Kleidung getragen werden (sog. Smart Clothing). Wearable Computers sind im Gegensatz zu den oben beschriebenen persönlichen Technologien immer operabel, nicht monopolisierend und erlauben freihändige Benutzung. Immer operabel heißt, dass Geräte nicht explizit aktiviert werden müssen, wie das bei Organizern und ähnlichen Geräten der Fall ist. Nicht monopolisierend heißt, dass die Benutzung nicht die ungeteilte Aufmerksamkeit des Anwenders verlangt und ihn so von anderen Tätigkeiten abhält. Mit diesen Eigenschaften ergeben sich eine Vielzahl praktischer Einsatzgebiete, etwa überall dort, wo Anwender Tätigkeiten durchführen, bei der sie keine Hand für die Bedienung unterstützender Informationstechnologien frei haben.

Die Idee, informationstechnische Unterstützung vom PC in die zu unterstützenden Anwendungen zurück zu verlagern findet noch stärker als in den beschriebenen persönlichen IAs ihren Ausdruck in der Schaffung von Informationsumgebungen. Hier steht im Gegensatz zu portablen Lösungen der Gedanke im Vordergrund, Informationstechnologie am Ort der Anwendung einzubetten und die reale Anwendungsumgebung als Schnittstelle zu Informationsdienstleistungen zu verwenden. Räume werden mit verschiedenen Information Appliances wie Tabs (elektronische Version von Post-Its), Pads (etwas größer und ersetzen Papier oder Bücher) und Boards (interaktive Tafeln) ausgestattet. Alle Geräte sind bewusst nicht personalisiert sondern als Artefakte der Arbeitsumgebung konzipiert. Neuere Arbeiten betten in ähnlicher Form verschiedene Information Appliances in Arbeitsumgebungen ein, gehen aber wesentlich weiter in der Verflechtung von Inneneinrichtung und Informationstechnologie.

Ein anderes Beispiel für mit der alltäglichen Umgebung verschmelzende IAs sind Ambiente-Medien, deren Anwendung zur Darstellung von Information in der räumlichen Umgebung am MIT MediaLab demonstriert wurde [Wisneski, Craig 1998]. Ambiente-Medien bringen Information durch die Manipulation von Artefakten der realen Umgebung zur Darstellung. Sie betten Informationsdarstellungen in das räumliche Ambiente in einer Form ein, die eine periphere und auch unterbewusste Wahrnehmung ermöglicht. So dargestellte Information kann beiläufig aufgenommen werden.

Die Vision der Information Appliances ist klar: Überwinde die Komplexität, erwirke die Einfachheit, mache es möglich, die Daten/Informationen zu kombinieren und gemeinsam zu nutzen, unabhängig von Gerät und Hersteller. Dabei sollen diese Geräte überall ihre Anwendungen finden, beim Sport, bei der Kunst, in der Industrie, bei der Kommunikation. Nötig dazu ist ein endanwender-orientiertes Denken der Entwickler und ein weltweites Abkommen über die entsprechende Infrastruktur, die den Austausch der Daten ermöglicht. Die erfolgreiche Familie der Information Appliances wird um die Menschen, die sie nutzen, und die Tätigkeiten, die sie unterstützen, herum gebaut.

3 Die Geräte der Zukunft

Wir befinden uns derzeit mitten in dem Trend "alles, immer, überall". Das Internet wird mobil zugreifbar, und auf dem Markt tauchen immer mehr persönliche "information appliances" wie drahtlos vernetzte PDAs, WAP-fähige Mobiltelefone oder elektronische Bücher und Reiseführer auf. Ermöglicht wird dies primär durch den weiter anhaltenden Fortschritt aller Zweige der Informationstechnik hin zum "kleiner, billiger, leistungsfähiger". Auch neue Entwicklungen der Materialwissenschaft (z.B. Miniatursensoren, "leuchtendes Plastik", "elektronische Tinte") und Fortschritte der Kommunikationstechnik (insbes. im drahtlosen Bereich) tragen dazu bei, dass es bald kleinste und spontan miteinander kommunizierende Rechner im Überfluss geben wird, die kaum mehr als solche wahrgenommen werden, da sie in Gebrauchsgegenstände eingebettet werden und so mit der alltäglichen Umgebung zu "smart objects" verschmelzen.

Information Appliances sind zurzeit für die Unternehmen noch kein großes Geschäft. Das liegt unter Anderem daran, dass sich die Computerindustrie noch nicht auf das Geschäftsmodell der Information Appliances, das ein Endnutzer-Modell ist, eingestellt hat. Die Industrie betreibt selbst immer noch das altbekannte Geschäftsmodell der Technologie-Produkte. Die Unternehmen müssen nun den Wechsel einleiten. In der High-Tech-Welt kommt der Profit

für das Unternehmen durch die Technologie selbst, es lebt vom Verkauf der Hard- und Software. Das Geschäftsmodell der Information Appliances ist vergleichbar mit dem von Konsumgütern, speziell mit dem von Haushaltsgeräten und Konsumelektronik. Der Gewinn wird hier durch das Anbieten des Content und sonstiger Dienstleistungen erbracht. Information Appliances sind Teil eines Ganzen und nicht als isolierte Geräte zu sehen.

Die Produkte der Informationstechnologie haben zu lange unter ihrer Technologie-Zentrierung gelitten. Nun ist die Zeit für den Wandel in Richtung einer endnutzer-orientierten Entwicklungsphilosophie gekommen. Menschen sind keine Maschinen, sie haben unterschiedliche Bedürfnisse; noch muss das Individuum zu der Technologie passen, es wird Zeit, dass die Technologie zum Individuum passt.

4 Literatur

- [Bergman, Eric 2000]: Information Appliances and Beyond. Academic Press, San Diego, 2000
- [Buse, Stephan 2002]: Der mobile Erfolg - Ergebnisse einer empirischen Untersuchung. In: Keuper, Frank (Hrsg.): Electronic Business und Mobile Business; Ansätze, Konzepte und Geschäftsmodelle. Wiesbaden: Gabler, 2002, S. 89-116
- [Keupner, Frank 2002]: Convergence-based View - ein strategie-strukturations-theoretischer Ansatz zum Management der Konvergenz digitaler Erlebniswelten. In: Keuper, Frank (Hrsg.): Electronic Business und Mobile Business; Ansätze, Konzepte und Geschäftsmodelle. Wiesbaden: Gabler, 2002, S. 603-654
- [Norman, Donald A. 1998]: The Invisible Computer. MIT Press, Cambridge, 1998
- [Schumann, Matthias 2002]: Betriebswirtschaftliche und technologische Grundlagen von E-Commerce und M-Commerce. In: Keuper, Frank (Hrsg.): Electronic Business und Mobile Business; Ansätze, Konzepte und Geschäftsmodelle. Wiesbaden: Gabler, 2002, S. 4-25
- [Weiser, Mark 1991]: The Computer for the 21st Century. Scientific American, September 1991, p. 66-75
- [Wisneski, Craig 1998]; Ishii, H.; Dahley, A.; Gorbett, M.; Brave, S.; Ullmer, B.; Yarin, P.: Ambient Display. Turning Architectural Space into an Interface between People and Digital Information. In: Proceedings of the First International Workshop on Cooperative Buildings (CoBuild'98), February 25-26, Darmstadt, Germany, LNCS Nr. 1370, Heidelberg, Springer-Verlag, 1998



Object Segmentation Using Colour for Security Applications on Mobile Devices

Maurice Danaher, Lennon Hopkins

Edith Cowan University
Perth, Western Australia 6050
m.danaher@ecu.edu.au, l.hopkins@ecu.edu.au

Abstract

This paper presents research-in-progress on a colour based method of object segmentation for use in biometric security methods on mobile devices. The limited processing capability of mobile devices has restricted the computational complexity of security algorithms that can be used. The purpose of our system is to use a colour based method to reduce the amount of information that must be processed by biometric algorithms. This is accomplished by creating a mug shot of a target individual from an image with an unknown background and unknown target/camera positions. Specific security algorithms can then be applied to the generated mug shot. An overall reduction in data processing will allow more efficient biometric security measures to be implemented.

1 Introduction

Our work is concerned with the development of a security application for preventing unauthorised access to mobile devices. We are researching an image recognition approach that uses colour to segment the image. Current methods of biometrics such as facial detection/recognition are based heavily on facial geometry and edge detection. Despite the accuracy of these methods, the amount of computational resources required for processing is still too great for mobile devices to process efficiently. In comparison to shape-based biometric methods, it has been suggested that colour be used as a method of identifying and segmenting possible target locations as it is far less computationally intensive [Wu et al 00]. Based on previous experimentation by researchers in the field of computer vision and image processing, the consensus



on colour-based methods of recognition is stated by Gonzalez and Woods (2002) “Colour is a powerful descriptor that often simplifies object identification and extraction from a scene.”

In this work a reduction in processing requirements is accomplished by restricting the area of the image on which computationally intensive processes, such as shape/edge based recognition algorithms, are applied. Current methods of facial/feature recognition are based on algorithms that make use of geometric shapes and must be applied to the entire image. The use of colour as a segmentation method has been based on the assumption that a “discontinuity in colour between two regions of an image implies different surfaces” [Maxwell 00].

2 Method

A mug shot is an image of a target object with controlled scale, lighting and background (which is usually a constant colour) [Bichsel & Pentland 94]. Control of these variables is not always possible in the natural environment. Our system generates a mug shot from an image with unknown camera/target positions and unknown lighting conditions. Biometric algorithms are then used to process the generated mug shots.

A region-growing algorithm is applied to the acquired image to determine which area(s) can be defined as skin colour. The image is then segmented according to the areas defined by the region-growing algorithm. These segments are then analysed to determine if holes exist. The holes correspond to the location of the eyes. These segments are then weighted according to their location from the centre of the image and output individually.

A number of system parameters must be set prior to image analysis. These parameters include setting the colour range that is defined as skin tone colours, the number of initial seed pixels to be used by the region-growing algorithm and the boundary parameters used for region segmentation. The parameter determining the number of initial seed pixels provides the foundation of the region-growing algorithm. Currently, experimentation with the number of pixel seeds is being conducted to determine the optimal number of seed values. If the seed value is too high the amount of overall processing required increases. Inversely, if the seed value is too low, the accuracy of the system decreases.

The colour threshold parameter determines the variance of pixels surrounding the seed pixels. This parameter has a significant impact on the performance of the region-growing algorithm. If the colour threshold parameter is too low, larger regions will be segmented increasing the processing required. Conversely, if the colour threshold value is too high, smaller (possibly incomplete) regions will be segmented and the accuracy of additional recognition/detection algorithms will decrease.

Following is a simplified version of the method that is used by our system (Figure 1):

- 1- A target image is acquired using a web camera at a high resolution (1024 x 768). Images can also be acquired using a digital camera.
- 2- Using a colour-based region-growing algorithm, the image is processed to determine which area of the image contains possible locations of a human head. A number of seed pixels are used to determine the colour regions.
- 3- Once the colour regions are determined and the extremity of the regions established, a border is defined to accommodate hair. The border is then used to segment the regions for further image processing.
- 4- To reduce the number of false positives identified by our system, an algorithm is applied that determines if there are holes within a colour region. If the system detects two holes in a single colour region (corresponding to the location of the eyes) the segment is designated as a potential head image. This is done as follows. An average colour for the segmented image is calculated. The colours of the pixels are compared to the average colour of the segment. If the colour deviance from the average is greater than a threshold value it is defined as a hole pixel. A ratio, based on the size of the segment, is calculated to determine the number of adjacent hole pixels that must be present in a segment. If the number of adjacent hole pixels is greater than or equal to the ratio value then the segment is defined as a potential head segment.
- 5- The potential head segments are weighted according to the distance they are located from the center of the image (it is assumed that the target is near the center of the image).
- 6- The head segments are then output as mug shots or passed on to other algorithms for facial/feature detection according to the weightings they have been assigned.

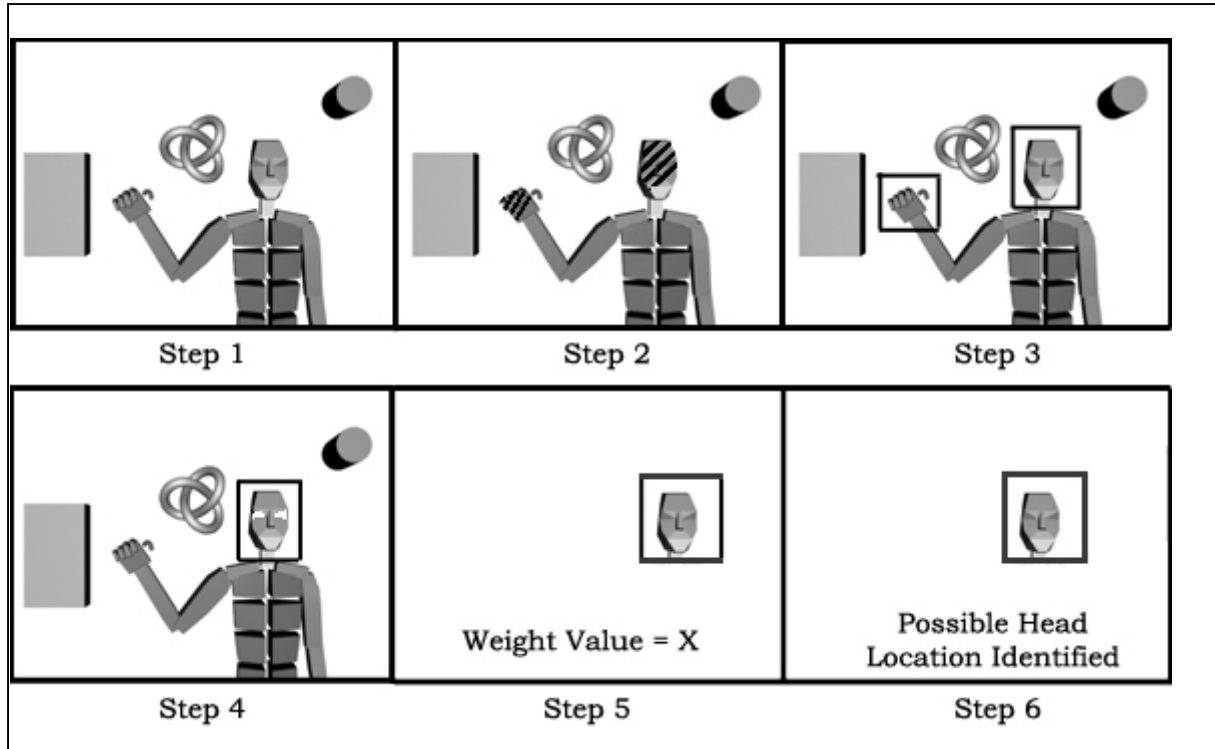


Figure 1: Graphical representation of the developed system.

3 Future Work

Preliminary experimentation with our system has revealed a number of flaws in the identification of potential head locations. The majority of flaws observed in our system are associated with background objects in the image and the parameters associated with them including colour threshold values and boundary values. Coloured illumination has also presented a problem for our system where non-skin coloured objects are illuminated with coloured lighting and exhibit colouring similar to skin tones. Our system has been developed with the assumption that the acquired images are taken with natural or white lighting.

The value of the colour variance threshold that is used by the region-growing algorithm becomes problematic if the background of the target is also skin-tone coloured or if it contains skin-tone coloured objects. Currently our system will identify an object of skin tone colour that contains holes regardless of the positioning of the identified holes. Further tests and experimentation with our system will include:

- Varying the range of colours that have been defined to be skin tone to reduce the number of false positives detected.

- Modifying the number of initial seed pixels to determine the optimum number required for maximum accuracy while keeping computation to a minimum.
- Modifying the algorithm for hole detection in a single colour region to prevent further processing of false positives e.g. if holes are aligned vertically disregard colour segment.
- Increasing the borders of the located target to prevent cropping an incomplete head region.
- Experimenting with the resolution of the initial image to minimize associated data processing.
- Experimenting with the threshold values of the average colour deviance within a segment.
- Experimenting with the ratio used to define the number of pixels that are part of a hole.

In conclusion the system we are developing supports the use of colour as its primary method of image segmentation. This approach may allow a greater variety of security applications, particularly methods based on biometric security, to be implemented on mobile devices.

4 References

- [Wu et al 00] Wu, Y. et al. "An Adaptive Self-Organizing Color Segmentation Algorithm with Application to Robust Real-Time Human Hand Localization", Intl. Proc. Asian Conf. On Computer Vision, 2000, Taiwan.
- [Gonzalez & Woods 02] Gonzalez, R.C. and R.E. Woods, Digital Image Processing, 2002, New Jersey: Prentice-Hall Inc.
- [Bichsel & Pentland 94] Bichsel, M. and A.P. Pentland, "Human Face Recognition and the Face Image Set's Topology", Image Understanding, 1994, 59(2): p. 254 - 261.
- [Maxwell 00] Maxwell, B.A., "Segmentation And Interpretation Of Multicolored Objects With Highlights". Computer Vision And Image Understanding, 2000, 77: p. 1 - 24.



Lerntheorie und kultureller Hintergrund - Einflussfaktoren bei der Gestaltung von Lernsystemen

Elisabeth Kamentz, Monika Schudnagis

Institut für Angewandte Sprachwissenschaften
Universität Hildesheim
Marienburger Platz 22
31141 Hildesheim
{ekam0028,schudnag}@rz.uni-hildesheim.de

Zusammenfassung

Um die Bedienfreundlichkeit der Benutzeroberflächen von Lernsystemen zu erhöhen, sollen Maßnahmen für deren Gestaltung formuliert werden, die auch die didaktische Basis des Systems einbeziehen. Daneben soll ein Benutzermodell entwickelt werden, das die Merkmale kulturspezifischen Lernverhaltens integriert und die Grundlage für eine Adaption der Benutzeroberfläche bildet. Zu diesem Zweck wird eine empirische Vorgehensweise gewählt, die sich am *rapid prototyping* Ansatz des Software-Engineering orientiert.

Abstract

In order to improve the usability of the user interface of learning systems, a model is to be developed which supports the process of designing user interfaces with regard to the didactic basis which will be used. Secondly a user model will be constructed which integrates cultural differences among learners and forms the basis for the adaptation of the user interface. In order to achieve these goals an empirical method is chosen which is related to the rapid prototyping approach in software engineering.

1 Einführung

Elektronische Lernumgebungen - seien es virtuelle, synchrone Lehrveranstaltungen oder individuell zu bearbeitende, asynchrone Lernprogramme - werden einerseits eingesetzt, um Mobilität zu vermeiden, können jedoch andererseits eine höhere Flexibilität und Mobilität ermöglichen. Geht es um die Entwicklung derartiger Lernsysteme, stehen zumeist inhaltliche oder technische



Probleme im Vordergrund. Fragen, die die Gestaltung der Benutzeroberfläche und damit die Mensch-Maschine-Interaktion betreffen, werden nur selten erörtert. Aus diesem Grund greift das Projekt SELIM (SoftwareErgonomie für Lernsysteme mit Multimedia) diesen Themenbereich auf und beleuchtet ihn vor dem Hintergrund verschiedener Faktoren, die die Gestaltung von Benutzeroberflächen derartiger Programme beeinflussen. Ein erstes Ziel besteht darin, Zusammenhänge zwischen der didaktischen Basis und dem softwareergonomischen Design multimedialer Lernsysteme aufzudecken und ein Modell für die Entwicklung derartiger Benutzeroberflächen zu erarbeiten. Den Ausgangspunkt dafür bildet die Tatsache, dass die Anwendung softwareergonomischer Gestaltungshilfen wie Styleguides, Richtlinien, etc. problembehaftet ist, da sie teilweise sehr spezifisch sind und zu Widersprüchen führen können bzw. nicht auf den konkreten Fall anwendbar sind. Im Gegenzug sind viele Hinweise zu allgemein, um wirklich als Hilfestellung dienen zu können. Alternativ dazu wäre ein Modell wünschenswert, das Gestaltungsprinzipien umfasst, die sich auf einen eingegrenzten Anwendungsbereich beziehen und damit einen gewissen Grad an Genauigkeit erreichen können ohne jedoch zu detailliert zu sein. Ein derartiges Modell liegt beispielsweise für den Bereich der Informationssysteme bereits vor (vgl. Krause & Womser-Hacker 1997) und könnte als Ansatzpunkt im Kontext der Lernsysteme dienen.

Das zweite Ziel des Projekts beinhaltet die Konstruktion einer Benutzermodellierungskomponente als Grundlage für die Adaptivität von Lernsystemen, die auf den unterschiedlichen Bedürfnissen Lernender mit variierendem kulturellen Hintergrund beruht. Dieser spielt bei der Herausbildung des Lernstils des Einzelnen eine wesentliche Rolle, nimmt Einfluss auf seine Erwartungen hinsichtlich der Gestaltung von interaktiven Lernsystemen und wirkt sich somit auf das gesamte Vorgehen während des Lernprozesses aus. Um den Lerngewohnheiten von Benutzern aus verschiedenen Kulturen (z.B. Austauschstudierenden) gerecht zu werden, sollen hier die kulturbedingten Unterschiede im Lernverhalten und in der Konzeption von interaktiven Lernsystemen (z.B. in den Bereichen Design, Navigationsstruktur, Didaktik) sowie die Merkmale der verschiedenen wissenschaftlichen Stile einzelner Kulturkreise berücksichtigt werden. Die Forschung auf dem Gebiet der Benutzermodellierung von interaktiven Lernsystemen hat sich bisher lediglich auf die Realisierung der Adaptivität im Bereich der Erfassung des individuellen Vorwissens, der Präferenzen und Ziele von Benutzern beschränkt (vgl. Kobsa & Wahlster 1989, Kobsa et al. 2000, Brusilovsky 1998). Angesichts der sprach- und kulturübergreifenden Kommunikation in der heutigen Welt und der damit verbundenen zunehmenden Studierendenmobilität besteht jedoch die Notwendigkeit der

Anpassung des Lernsystemverhaltens an die gewohnten Denkmuster und Lernstile von Studierenden aus unterschiedlichen Kulturen.

2 Theoretische Grundlagen

2.1 Lerntheorien und ihre programmtechnische Umsetzung

Es existiert eine Reihe von Lerntheorien, die ein breites Spektrum an Auffassungen darüber abdecken, wie Lernen vor sich geht und wie es gefördert werden kann (vgl. die Ausführungen zu Behaviorismus, Kognitivismus und Konstruktivismus in Blumenstengel 1998). Interessant ist in diesem Zusammenhang, wie sich einzelne Theorien programmtechnisch in Form von Lernsystemen umsetzen lassen. Diese reichen von behavioristisch motivierten drill & practice Programmen über kognitivistisch orientierte, adaptive Systeme (insbesondere Intelligente Tutorielle Systeme (ITS), vgl. Blumenstengel 1998:12), bis zu konstruktivistischen Lernumgebungen, die als „Informations- und Werkzeugangebot für selbstgesteuerte Lernprozesse“ (Tulodziecki et al. 1996:47) ein individuelles Vorgehen fördern sollen. Offensichtlich variieren Systemstruktur und Benutzeroberfläche stark in Abhängigkeit davon, welche Lerntheorie einem Programm zugrunde liegt: Während in behavioristischen Systemen die Mensch-Maschine-Interaktion im Wesentlichen durch das System bestimmt wird, gesteht ein konstruktivistisches Konzept dem Benutzer wesentlich mehr Entscheidungsfreiheit zu.

Jede Lerntheorie betont bestimmte Aspekte des Lernens und weist entsprechende Vor- und Nachteile auf. So fördert der Behaviorismus den unhinterfragten Erwerb von Grundwissen oder Basisfähigkeiten (Baumgartner & Payr 1997), behavioristisch orientierte Lernsysteme werden jedoch häufig nach kurzer Zeit als langweilig empfunden. Die kognitivistische Theorie wird im Bereich der Problemlösung als besonders geeignet betrachtet, es erweist sich jedoch als schwierig, Erkenntnisse über interne Vorgänge beim Lernenden in nicht-experimentellen Situationen zu erlangen (Kerres 1998:57). Dies ist jedoch unentbehrlich, um das Wissen so aufzubereiten, dass der Aufbau eines korrekten mentalen Modells beim Lernen ermöglicht wird. Die konstruktivistische Lerntheorie fördert insbesondere das Bewältigen komplexer Situationen, wobei Systeme, die sich auf diese Sichtweise gründen, oftmals überfordern, da der Lernende den Lernprozess selbstgesteuert und aktiv gestalten muss (Blumenstengel 1998:125). Es ist daher sicherlich nicht richtig, eine Theorie als die einzig wahre zu bezeichnen, da jede Sichtweise sehr unterschiedliche Schwerpunkte setzt und davon auszugehen ist, dass es Situationen gibt, in denen sich eine bestimmte Theorie besonders gut als Basis für die

Entwicklung eines Lernsystems eignet. Die Schwächen der einzelnen Ansätze lassen sich möglicherweise durch die Kombination von Elementen aus verschiedenen Theorien kompensieren. So soll dem Problem der Überforderung der Lernenden, das bei konstruktivistisch geprägten Lernumgebungen aufgrund der fehlenden Anleitung im selbstgesteuerten Lernprozess auftritt, beispielsweise durch einen integrativen Ansatz entgegengewirkt werden, der das instruktionalistisch motivierte Element der (optionalen) Führung in die konstruktivistische Sichtweise einbindet (Reinmann-Rothmeier & Mandl 2001). Im Rahmen von SELIM soll geprüft werden, inwieweit eine derartige Kombination von Merkmalen sinnvoll sein kann und inwiefern sie sich programmtechnisch realisieren lässt.

2.2 Kulturbedingte Unterschiede im wissenschaftlichen Stil und Lernverhalten

Das Gebiet der interkulturellen Kommunikationsforschung umfasst einen sehr weiten Bereich (vgl. Hofstede 1993, Hall/Hall 1990, Beneke 2001). Im Zusammenhang mit der Entwicklung von Lernsystemen sind nicht nur die allgemeinen Besonderheiten interessant, durch die sich verschiedene Gesellschaften auszeichnen, sondern vor allem die aus ihnen resultierenden kulturell bedingten Unterschiede im wissenschaftlichen Stil, insbesondere im Hinblick auf Diskursstrukturen, sowie die Auswirkungen von Kultur in Lehr-/ Lernkontexten. Hier ist neben dem kognitiven Aspekt des Lernstils auch das soziale Verhältnis zwischen Lehrenden und Lernenden relevant.

In seiner kulturvergleichenden Betrachtung der Wissenschaften unterscheidet der norwegische Sozialwissenschaftler Johan Galtung (vgl. Galtung 1981) verschiedene kulturspezifische wissenschaftliche Stile, mit denen sich die Unterschiede in den Diskursstrukturen verschiedener Kulturen erklären lassen. Er differenziert zwischen drei westlichen und einem östlichen Kulturkreis, in denen sich die Wissenschaft durch relativ homogene Methoden der Wissensvermittlung und -darstellung auszeichnet, z.B. in Bezug auf die Art der Informationsstrukturierung, die Informationstiefe oder die Sequenzierung der Inhaltsbausteine. Er nimmt dabei die folgende Zuordnung einzelner Länder zu den definierten intellektuellen Stilen vor, die sich aus den wissenschaftlichen Kulturen Großbritanniens bzw. der USA, Deutschlands, Frankreichs und Japans entwickelt haben: 1) sachsonischer intellektueller Stil (Länder des Commonwealth, USA), 2) teutonischer intellektueller Stil (deutschsprachige Länder, Osteuropa, Russland), 3) gallischer intellektueller Stil (Frankreich, Italien, Spanien, Portugal, Südamerika) und 4) nipponischer intellektueller Stil (Japan und fernöstliche Länder, Naher Osten).

Um Profile der wissenschaftlichen Stile für die einzelnen Ländergruppen erstellen zu können, hat Galtung in seiner Arbeit vier Dimensionen festgelegt, die in allen wissenschaftlichen Stilen in unterschiedlicher Ausprägung vorhanden sind: 1) Paradigmenanalyse, 2) Thesenproduktion, 3) Theoriebildung, 4) Kommentar über andere Intellektuelle bzgl. der ersten drei Kategorien. Galtung verankert seine Typologie der intellektuellen Stile in der gesellschaftlichen Struktur einzelner Kulturen. Wir gehen davon aus, dass sich diese u.a. auch auf die Konzeption von Lehrmaterial und die Gestaltung von Lernprozessen auswirkt und Galtungs Einteilung daher auch als komplementär zum kulturspezifischen Lernstil angesehen werden kann.

Auf der Grundlage der im Rahmen seiner interkulturellen Diskursforschung durchgeführten Analyse von expositorischen Texten Studierender aus verschiedenen Kulturen sowie der Regeln wissenschaftlichen Schreibens identifiziert Clyne (vgl. Clyne 1991, 1994) Unterschiede in der Absatzstruktur und der Pragmatik der schriftlichen wissenschaftlichen Diskurse. Er definiert mehrere Kategorien, in denen die Unterschiede zwischen deutschen und englischen Diskursstrukturen besonders sichtbar werden:

- Linearität vs. Digressivität
- Formorientierung vs. Inhaltsorientierung
- Abstraktion vs. Konkretion der Inhalte
- writer responsibility vs. reader responsibility
- Einsatz von Advance Organizern
- Integration von Datenmaterial

Die hier ansatzweise vorgestellten Unterschiede in den unterschiedlichen wissenschaftlichen Kulturen lassen sich mit Hilfe der vier Kulturdimensionen nach Hofstede (Hofstede 1993) analysieren und erklären. Insbesondere die Dimensionen 'Individualismus vs. Kollektivismus', 'Unsicherheitsvermeidung' und 'Machtdistanz' haben den akademischen Stil einzelner Kulturkreise beeinflusst. Im Folgenden werden die wesentlichen Auswirkungen dieser drei Kulturdimensionen auf die Gestaltung von Lernsituationen dargestellt (vgl. hierzu Hofstede 1986).

1. Individualismus vs. Kollektivismus: Diese Dimension beschreibt, inwieweit sich der Mensch als Individuum bzw. als Teil einer Gemeinschaft begreift. Im Hinblick auf die Gestaltung von Lernsituation ergeben sich zwischen individualistischen und kollektivistischen Kulturen vor allem Unterschiede hinsichtlich des Lernzwecks. Für einen individualistisch orientierten Lerner besteht der Zweck des Lernens darin, zu erfahren, wie man Neues erlernen kann, um mit unbekannten, unerwarteten Situationen umgehen zu können. Man geht von der Notwendigkeit des lebenslangen Lernens aus. Dagegen ist der Lernzweck für einen kollektivistisch geprägten

Schüler eher im Beherrschen von vorgegebenen Fakten und Fertigkeiten sowie der Anpassung an gesellschaftliche Tugenden und Traditionen zu sehen. Man lernt, wie etwas gemacht wird, um von der sozialen Gruppe, der man angehört, akzeptiert zu werden. Darüber hinaus sind Unterschiede im Diskussionsverhalten festzustellen. In kollektivistischen Kulturen werden Schüler erst sprechen, wenn sie von ihrem Lehrer persönlich dazu aufgefordert wurden. Individualistisch orientierte Kulturen dagegen legen Wert auf offene Konfrontationen, bei denen eine Vielzahl gegenteiliger Meinungen eine Selbstverständlichkeit darstellt. Zu dem individualistisch geprägten Kulturkreis gehören die USA und die Länder des Commonwealth sowie Nordeuropa, während die meisten asiatischen Länder, arabische und lateinamerikanische Länder (mit Ausnahme Brasiliens) eine kollektivistische Orientierung aufweisen (vgl. Beneke 2001).

2. Machtdistanz: Machtdistanz betrifft das Ausmaß, bis zu welchem die weniger mächtigen Mitglieder einer Gesellschaft eine ungleiche Machtverteilung erwarten und akzeptieren. In bezug auf Lernsituationen ergibt sich für den Lehrenden in Kulturen mit niedriger Machtdistanz, wie z.B. Deutschland, Skandinavien, USA, Kanada, Australien, 'lediglich' die Position eines Fachexperten (Primus inter Pares), der losgelöstes (neutrales) Wissen vermittelt und von seinen Schülern Eigeninitiative erwartet. In Kulturen mit großer Machtdistanz dagegen, zu denen arabische Länder, Lateinamerika (mit Ausnahme Argentinien) sowie Ost- und Südostasien; Frankreich, Spanien, Belgien, Türkei und eingeschränkt auch Italien zählen (vgl. Beneke 2001), geht jegliche Initiative vom Lehrenden aus, der als 'Guru' sein eigenes Wissen vermittelt und als Autoritätsperson entsprechenden Respekt erwartet.
3. Unsicherheitsvermeidung: Unsicherheitsvermeidung drückt eine Strukturiertheitserwartung aus, d.h. Unbekanntes, Unerwartetes und Vages soll in allen Lebensbereichen möglichst vermieden oder ausgeschlossen werden. Kulturen mit niedriger Unsicherheitsvermeidung (z.B. Großbritannien, USA, Kanada, Australien, südostasiatische Länder), die sich durch Risikobereitschaft und die Akzeptanz von Ambiguität und Neuem als einer normalen Erscheinung im Leben auszeichnen, weisen im Lernkontext eine Vorliebe für kontroverse Diskussionen und Open-End-Lernsituationen auf. Länder mit einer hohen Unsicherheitsvermeidung, wie z.B. lateinamerikanische und islamische Länder, Mittelmeerländer, deutschsprachige Länder sowie einige ostasiatische Länder wie Japan oder Südkorea, bevorzugen dagegen klar strukturierte Lernsituationen und die Vorgabe von korrekten Antworten.

An dieser Stelle ist nun die Frage interessant, wie sich die Einflüsse der Kulturdimensionen und der zuvor vorgestellten unterschiedlichen wissenschaftlichen Stile auf softwareergonomische Aspekte wie Farbgebung, Bildschirmorganisation, Interaktionsmöglichkeiten und Navigationsstruktur sowie die Präsentation der Lehrinhalte und das didaktische Konzept von Lernprogrammen aus verschiedenen Kulturen auswirken. Mit Hilfe der Evaluation anhand eines auf der Grundlage der hier vorgestellten Erkenntnisse im Bereich der interkulturellen Kommunikationsforschung entwickelten Kriterienkatalogs soll dies im Rahmen unserer Studie festgestellt werden (siehe Abschnitte 4.2 und 5.2).

3 Angestrebte Ziele

Sowohl softwareergonomische Gestaltungshilfen als auch Maßnahmen, die auf Anpassungsleistungen basieren, haben den Zweck, die Arbeit mit interaktiven Systemen zu erleichtern. Im Kontext der Lernsysteme bedeutet das, dass sich im Zuge der Verringerung von Interaktionsproblemen und der stärkeren Orientierung am Benutzer auch die Effektivität beim Lernen steigern lässt.

3.1 Entwicklung eines Gestaltungsmodells

Da existierende Gestaltungshilfen bei der Entwicklung spezifischer Benutzeroberflächen häufig nicht entscheidend weiterhelfen, soll für einen eingegrenzten Anwendungsfall (hier: Lernsysteme mit dem Lernziel Übung) ein Modell entwickelt werden, das verschiedene Maßnahmen umfasst, die die Entwicklung derartiger Systeme fördern. Besonderes Augenmerk soll dabei auf die lerntheoretische Basis gelegt werden. Die Auswahl einer Lerntheorie orientiert sich an den Faktoren Lerninhalt (content), Lernereigenschaften und Lernziel (z.B. Auswendiglernen von Fakten, Problemlösungsfähigkeit erlangen, kritische Auseinandersetzung mit einem Thema). Dabei ist zu beachten, dass die Wünsche der Lernenden sehr unterschiedlich sein können (z.B. Inhalte lieber selber erarbeiten oder lieber präsentiert bekommen) und dass sich die Anforderungen im Laufe des Lernprozesses auch verändern können (z.B. vom Anfänger auf einem Gebiet zum Fortgeschrittenen). Das bedeutet, dass nicht nur innerhalb der Zielgruppe Variationen vorliegen, sondern dass auch der einzelne Lernende einer dynamischen Entwicklung unterliegt, die beim Systementwurf zu berücksichtigen ist. Idealerweise beinhaltet das zu entwickelnde Modell Gestaltungsmaßnahmen, die den Entwurf von Lernsystemen dergestalt unterstützen, dass es dem Lernenden (je nach Wissensstand und Vorlieben) unterschiedliche Sichten auf das System ermöglicht, damit er – auch bei sich verändernden Anforderungen – innerhalb des Programms verbleiben

kann. Ein ähnliches Konzept wurde bereits für den Bereich der Informationssysteme in Form der strikten Objektorientierung vorgeschlagen (vgl. Krause & Womser-Hacker 1997).

3.2 Entwicklung einer Benutzermodellierungskomponente

Ein weiteres Ziel auf dem Weg, Lernsysteme benutzerorientiert zu gestalten, ist die Entwicklung einer Benutzermodellierungskomponente, die dem System eine Anpassung an die Präferenzen, Erwartungen und Lernstile von Benutzern aus verschiedenen Kulturen ermöglicht. Hierzu ist zunächst die Ermittlung von Merkmalen notwendig, die für Lernprogramme aus verschiedenen Kulturen charakteristisch sind, insbesondere in den Bereichen Design, Interaktions- und Navigationsmöglichkeiten, Inhaltspräsentation und Didaktik. Ergänzend werden mit Hilfe einer kulturvergleichenden Lernstilanalyse (vgl. hierzu auch Barmeyer 2000) die Merkmale des Lernverhaltes von Studierenden mit unterschiedlichem kulturellen Hintergrund erforscht. Anschließend sollen die erarbeiteten Profile des Aufbaus von Lernprogrammen sowie die ermittelten Lernstilmerkmale Studierender aus ausgewählten Kulturen eine Basis für die Konzeption der Benutzermodellierungskomponente des adaptiven Lernsystems für Anwender mit unterschiedlichen kulturellen Backgrounds bilden.

4 Methode und Vorgehensweise

Die Methode, die in beiden Teilen unserer Studie im Zusammenhang mit der Evaluation der entwickelten Lernsysteme zum Einsatz kommt, ist das Rapid-Prototyping-Verfahren. Im Sinne der formativen Evaluation sieht Rapid Prototyping die Einbeziehung potentieller Benutzer in den Systementwicklungsprozess zu einem möglichst frühen Zeitpunkt vor. Ihre Partizipation ermöglicht das Testen der Systementwürfe in realistischen Situationen und dient somit deren iterativer Verbesserung und der Erhöhung der Benutzerakzeptanz. Die entwickelten Prototypen werden hinsichtlich der Benutzerorientierung, des Nutzungsverhaltens und des Grades des Lernerfolgs getestet, wobei die Versuchspersonen zum ‚lauten Denken‘ animiert werden. Die Protokollierung der Tests erfolgt mittels Videoaufzeichnung. Der Entwicklungsprozess wird mehrere Testzyklen umfassen, deren Ergebnisse wiederum in die Entwicklung mit einfließen sollen.

4.1 Oberflächengestaltung: Prototyp-Entwicklung und Benutzertest

Um ein Modell für das Benutzeroberflächendesign von Lernsystemen bilden zu können, wurden zunächst zwei Systemprototypen konzipiert, die sich auf unterschiedliche lerntheoretische Weise an eine Thematik annähern. Dazu musste aufgrund der empirisch orientierten Vorgehensweise zunächst ein Gegenstandsbereich ausgewählt werden; dieser sollte aus der informationswissenschaftlichen Hochschullehre stammen. Da einige Lehrformen (Vorlesung, Seminar und Projekt) bereits im Rahmen des Projekts ‚Virtueller Campus‘ (vgl. Caroli 2000) abgedeckt wurden, fiel die Wahl auf eine Lerneinheit aus der die Vorlesung ‚Einführung in die Informationswissenschaft‘ begleitenden Übung. Diese Übung scheint aufgrund ihres interaktiven Charakters besonders geeignet, da der Lernende nicht nur rezipiert, sondern auch die Möglichkeit bekommt, aktiv zu handeln. Als inhaltliche Basis der zu entwickelnden Lernsystem-Prototypen sollte der Themenbereich ‚Statistische Evaluierung von Information Retrieval Systemen‘ umgesetzt werden. Bei der Erarbeitung der lerntheoretischen Grundlagen zeigte sich, dass es aufgrund der fließenden Übergänge nicht möglich und aufgrund der Theorieschwächen oftmals nicht wünschenswert ist, Systeme strikt nach einer einzigen Auffassung vom Lernen zu konzipieren (vgl. integrativer Ansatz in Abschnitt 2.1). Daher wurden zwei Systeme entwickelt, die verschiedene Ansätze miteinander kombinieren. Das System *bekog* verfolgt zum einen das behavioristische Prinzip, das sich in einer relativ linearen Systemstruktur manifestiert, bei der Aufgaben auf die entsprechenden Lerneinheiten folgen. Dieser Prototyp ist angereichert mit kognitionspsychologisch motivierten Elementen, die die Aufmerksamkeit steigern und den Aufbau mentaler Modelle fördern sollen (angelehnt an Ruppell & Pflöging 2000). Das System *kogkons* vereint konstruktivistische und kognitivistische Elemente. Es besteht aus zwei Bereichen (Thema und Arbeitsbereich), die sich optisch stark voneinander unterscheiden, und verfügt über weitaus reichere Navigationsmöglichkeiten, die die freie Navigation durch das Angebot eines idealen Pfades ergänzen. Der Schwerpunkt liegt hier auf der Erarbeitung von Wissen, daher ist der Bereich der Aufgabenbearbeitung zentral. Aufmerksamkeit soll in erster Linie dadurch geweckt werden, dass ein authentisches Problem als Ausgangspunkt präsentiert und durchgehend bearbeitet wird. Als kognitivistisch orientiert ist hier das Zusammenspiel von gesprochener Sprache und Animation zu betrachten, das der begrenzten Verarbeitungskapazität des menschlichen Gedächtnisses entgegenwirken soll.

Diese beiden Systeme sollten im Rahmen eines Benutzertests in Hinblick auf ihre Benutzbarkeit, die Akzeptanz bei den Benutzern, aber auch auf ihre Lern-

wirksamkeit hin miteinander verglichen werden. Das Ziel ist dabei, Erkenntnisse darüber zu gewinnen, an welchen Stellen

- Entscheidungen in der Gestaltung der Benutzeroberfläche bestätigt werden bzw. nochmals zu überdenken sind
- eine stärkere Anpassung an die Benutzerbedürfnisse erforderlich ist
- die Stärken und Schwächen der beiden Grundkonzepte liegen
- sich Möglichkeiten der Integration beider Ansätze zu einem komplexen Konzept bieten.

An dem Benutzertest nahmen 17 Studierende der Informationswissenschaft im ersten Semester teil, die die Systeme als Alternative zur Übungsstunde verwendeten (5 Versuche pro System). Da im Rahmen des Experiments die Interaktion mit anderen Lernenden nur schwer nachzustellen ist, wurde - so weit möglich - in Zweiergruppen gearbeitet und auf diese Weise die Einbindung der Lernenden in einen sozialen Kontext als Merkmal konstruktivistisch orientierter Lernumgebungen umgesetzt. Diese Konstellation sollte darüber hinaus das „laute Denken“ bei den Versuchspersonen fördern, da das Ansprechen eines direkten Gegenübers eine natürlichere Situation darstellt als das Führen von Selbstgesprächen. Die Sitzungen wurden per Video protokolliert; zusätzlich füllten die Versuchspersonen vor und nach dem Experiment Fragebögen aus, mit deren Hilfe Computerkenntnisse, eine Beurteilung des jeweiligen Systems und das Vorher-/Nachher-Wissen der Versuchspersonen über das Thema festgehalten werden sollten. Nach der Transkribierung der Videoprotokolle erfolgte deren Auswertung im Hinblick auf Navigationsverhalten, Bedienprobleme, Probleme bei der Aufgabenlösung, etc. Auf der Basis der dabei erzielten Ergebnisse werden die Prototypen überarbeitet und erste Folgerungen bzgl. der Modellbildung abgeleitet.

4.2 Evaluation von Lernprogrammen/Webtutorials

Die Evaluation von Lernprogrammen auf CD-ROM und Webtutorials aus verschiedenen Kulturen hinsichtlich ihres didaktischen Aufbaus, softwareergonomischer Aspekte und der Merkmale der inhaltlichen Informationsdarstellung wird anhand eines umfassenden Katalogs von mehr als 50 Kriterien durchgeführt. Der Kriterienkatalog gliedert sich in die folgenden Bereiche:

1. Allgemeine Angaben zum Programm (z.B. Quelle, Themenbereich, Kultur und Sprache des Autors)
2. Softwareergonomie:
 - Design (z.B. Einsatz von Farben und multimedialen Elementen; Verwendung von Metaphern, Icons, Typografie)

- Interaktion und Navigation (z.B. Realisierung von Menüs; Einsatz von verschiedenen Navigationstools)
- 3. Inhalt (z.B. Strukturierung der Information; Art, Zusammensetzung und Sequenzierung von Inhaltsbausteinen wie Theorie, Fakten, Beispiele; Einsatz von Sprache)
- 4. Didaktik (z.B. allgemeine didaktische Konzeption, Lernzielarten, Einsatz von unterschiedlichen Aufgabentypen, Gestaltung des Feedback)

Das Ziel dieser Evaluation ist es herauszufinden, in welchen Bereichen des Programmaufbaus sich kulturspezifische Merkmale feststellen lassen und um welche Merkmale es sich handelt. Da bei jeder Evaluation ein anderes Ziel verfolgt wird (z.B. Bewertung der Interaktion und des didaktischen Konzepts im Hinblick auf verschiedene Zielgruppen), muss für jede einzelne Untersuchung, von allgemeinen Bewertungsmaßstäben ausgehend, eine spezifische Kriterienliste erarbeitet werden, bei der Schwerpunkte gesetzt werden, die den angestrebten Zweck erreichen lassen. Der im Rahmen dieser Arbeit entwickelte Katalog enthält Kriterien, die es erlauben, die kulturell bedingten Merkmale der bewerteten Programme zu erfassen, jedoch ohne deren Güte im Hinblick auf Usability oder Didaktik beurteilen zu wollen.

4.3 Befragung zu Lernstilen und Computernutzung

Die Befragung soll der Ergänzung der Evaluationsergebnisse im Hinblick auf die Entwicklung der Benutzermodellierungskomponente des Lernsystems dienen. In unserer Arbeit gehen wir davon aus, dass der kognitive Stil, d.h. das Denken sowie die Arbeits-, Problemlöse- und Lerntechniken des Einzelnen nicht nur durch seine individuelle Veranlagung, sondern auch durch kulturelle Einflüsse gebildet werden und daher zu dem Spektrum kultureller Faktoren gehören, die bei der Konzeption und Entwicklung von benutzerorientierten Lernsystemen Berücksichtigung finden müssen. Insbesondere für das Lernen mit hypermedialen Lernsystemen, die wegen flexibler Darstellungs- und Navigationsmöglichkeiten eine Anpassung an unterschiedliche Lernertypen erlauben (content-level adaptation, link-level adaptation, vgl. Brusilovsky 1998), ist die Lernstilforschung von besonderer Relevanz.

Die Durchführung einer kulturvergleichenden Lernstilanalyse sowie der Analyse der Computernutzung ist im Rahmen einer Umfrage unter potentiellen Systemnutzern aus den ausgewählten Kulturen erfolgt. Der Fragebogen basiert auf der Idee des Learning Style Inventory (LSI), das von dem amerikanischen Psychologen David A. Kolb entwickelt wurde und zu den bekanntesten Untersuchungsinstrumenten zur Operationalisierung von individuellen

Lernstilen zählt. Es basiert auf der Theorie des Experiential Learning, in deren Zentrum die Beschreibung des Lernprozesses steht (vgl. Kolb 1984). Dieser wird als ein vierphasiger Zyklus begriffen: die konkrete Erfahrung (Concrete Experience, CE) bildet eine Basis für reflektierende Betrachtung (Reflective Observation, RO), die Beobachtungen werden von dem Lernenden bei der Entwicklung von Theorien und Modellen verwendet (Abstract Conceptualization, AC), aus denen wiederum Hypothesen abgeleitet werden können, die durch praktische Erfahrung bestätigt werden sollen (Active Experimentation, AE). Diese Phasen bilden die gegensätzlichen Ausprägungen von zwei Dimensionen: konkrete vs. abstrakte Erfassung von Informationen und aktive vs. reflektierende Verarbeitung der erfassten Informationen. Auf der Grundlage von empirischen Untersuchungen reduzierte Kolb verschiedene Lernverhalten auf vier Lernstile, die sich im Verlauf des Lernprozesses abwechseln (vgl. Kolb et al. 1995):

1. Accommodators (Pragmatiker) lernen hauptsächlich aus praktischer Erfahrung, sind risikobereit und aktiv, flexibel und experimentierfreudig. Sie ändern zuvor aufgestellte Pläne, wenn es die Situation erfordert, und tendieren mehr zum instinktiven und spontanen Handeln als zu logischer Analyse.
2. Divergers (Universalisten) beobachten Situationen eher, als zu handeln, und versuchen, die Konsequenzen unbekannter Situationen vorherzusehen. Sie betrachten konkrete Sachverhalte aus verschiedenen Perspektiven, sind jedoch keine Theoretiker, die einer abstrakten Logik folgen. Sie sammeln innovative Ideen (Brainstorming) und zeichnen sich durch Kreativität aus.
3. Convergers (Spezialisten) wenden Theorien und Modelle praktisch an, stellen Hypothesen auf, gehen dabei pragmatisch aber eher unsystematisch vor. Ihre Stärken liegen in der Fähigkeit, neue Lösungsstrategien anzuwenden, dabei sind sie aufgaben- und zielorientiert und bevorzugen Aufgaben mit einer richtigen Lösung.
4. Assimilators (Theoretiker) konzentrieren sich auf die Entwicklung von Theorien und Modellen durch induktives Schlussfolgern, wobei sie analytisch und systematisch vorgehen. Sie fassen große Informationsmengen zu einem logischen Ganzen zusammen und betrachten Situationen aus vielen unterschiedlichen Perspektiven. Die logische Stichhaltigkeit einer Theorie ist ihnen häufig wichtiger als ihr praktischer Wert.

Wir betrachten diese Typologie als geeignet, um kulturbedingte Lernstilunterschiede zu erforschen, da sich die in dieser Einteilung definierten Lernstilmerkmale mit den kulturbedingten Unterschieden im wissenschaftlichen Stil und in Lehr-/Lernsituationen (siehe Abschnitt 2.2) in Verbindung bringen las-

sen. Eine weitere Kategorisierung zur Erfassung der für das Lernen relevanten Persönlichkeitsmerkmale, die hier nicht unerwähnt bleiben soll, jedoch im Zusammenhang mit der Kulturspezifität des Lernens eine geringere Rolle spielen dürfte, ist die Differenzierung zwischen Lernstilen und Kognitionsstilen (vgl. Jonassen/Grabowski 1993). Während unter Lernstilen (im engeren Sinne) selbst gesetzte Präferenzen der Lernenden für unterschiedliche Lehr- und Lernarten verstanden werden (Modell von Kolb), stellen Kognitionsstile individuelle Persönlichkeitsmerkmale dar, die die Herangehensweisen des Lernenden bei der Informationssammlung und -organisation beeinflussen. An dieser Stelle wird eine weitere Unterscheidung vorgenommen: Bei der *Informationssammlung* steht die Präferenz für die Wahrnehmung durch einen der Sinne im Vordergrund – dementsprechend wird zwischen dem visuellen, auditiven und kinästhetischen Lerntyp differenziert (Sinnesmodalität), wobei Kombinationsformen möglich sind. Die Präferenz für bildliche vs. geschriebene oder gesprochene verbale Informationsaufnahme (Codalität) stellt eine weitere Unterteilung dar. Im Zusammenhang mit der *Informationsorganisation* steht der Grad der Detailorientierung im Vordergrund – hier unterscheidet man zwischen Serialisten und Holisten. Die Serialisten zeichnen sich durch ausgeprägte Detailorientierung und eine lineare Vorgehensweise aus, die allerdings dazu führt, dass ein übergeordnetes Gesamtkonzept erst später aufgebaut wird (bottom-up). Dagegen betrachten Holisten mehrere Aspekte gleichzeitig und stellen zuerst einen globalen Gesamtzusammenhang mit allen relevanten Querverbindungen her (top-down), bevor sie sich mit Details beschäftigen (vgl. Blumenstengel 1998).

Der zweite Teil der Befragung soll Erkenntnisse darüber liefern, ob in den verschiedenen Kulturen unterschiedliche Zugangsweisen zu Computern vorherrschen. Die Fragen beziehen sich u.a. auf Computerkenntnisse und Nutzungsgewohnheiten, Einstellungen zur Computertechnologie, Erfahrungen mit Computerkursen sowie Präferenzen bzgl. der funktionellen und inhaltlichen Gestaltung von Lernprogrammen.

5 Erste Ergebnisse

Erste Ergebnisse liegen sowohl seitens der Benutzertests als auch in bezug auf die Analyse kultureller Aspekte für die Konstruktion einer Benutzermodellierungskomponente vor.

5.1 Ergebnisse des Benutzertests

Aus der Analyse der Fragebögen bzw. Videoprotokolle konnten einige Erkenntnisse als Basis für eine Überarbeitung der Prototypen innerhalb des nächsten Prototyping-Zyklus gewonnen werden.

Layout: Die optische Gestaltung des Systems stellt einen sehr wichtigen Aspekt für die Versuchspersonen dar, den sie auch von sich aus thematisierten. Der großzügige Einsatz von Farben, der eine positive Grundstimmung vermitteln soll, gleichzeitig aber auch in der Verwendung eines Farbcodes als Orientierungshilfe begründet ist, wird von einem Teil der Versuchspersonen begrüßt, der andere Teil würde eine nüchternere Gestaltung vorziehen. Daher erscheint es angebracht, den Benutzern eine Individualisierung der Oberfläche in Hinblick auf die optische Gestaltung zu ermöglichen.

Navigation: Die umfangreichen Navigationsmöglichkeiten, die im System *kogkons* verfügbar sind, werden kaum genutzt, weil sie den im Umgang mit dem System unerfahrenen Benutzer womöglich überfordern. Stattdessen wird der vorgegebene Pfad von den meisten Versuchspersonen bevorzugt. Dagegen empfinden einige Testteilnehmer die Möglichkeiten zur Navigation im System *bekog* als zu restringiert und gelegentlich als suboptimal. Ideal wäre ein variables Angebot an Navigationsoptionen, das an den Einsatzzweck des Systems bzw. die Anforderungen des Benutzers angepasst werden kann.

Texte: Die Beurteilung des Umfangs der Texte ist sehr konträr: während zwei Drittel der Versuchspersonen die Texte positiv bewerten, halten einige sie für zu ausführlich. Andere bemängeln wiederum deren Kürze. Dies spricht für eine gestufte Informationsdarstellung, die sich aus einer knappen Darstellung und weiterführender Information mit Beispielen zusammensetzen könnte. Auffallend war auch, dass sich einige Versuchspersonen gegen ähnlich lautende Texte aussprachen, da dies zu Orientierungsproblemen führen kann.

Hilfe und Feedback: Die Unterstützung bei der Aufgabenlösung sollte ausgebaut und das Feedback detaillierter gestaltet werden, da einige Versuchspersonen beispielsweise die Darstellung von Rechenwegen wünschten. Denkbar wäre auch die Integration einer Komponente, die Fehler genauer analysiert und an die Situation angepasstes Feedback präsentiert. Bemerkenswert ist auch die Tatsache, dass die Meldungen oft sofort wieder weggeklickt werden, wenn sie zu häufig auftreten, so dass möglicherweise eine Variation in deren optischer Aufbereitung vorzunehmen ist, um die Aufmerksamkeit der Benutzer aufrechtzuerhalten.

Lernerfolg: Bei der Ermittlung des Lernerfolgs wurden vor und nach der Arbeit mit dem System Fragen zum vorliegenden Thema gestellt und die Antworten dazu verglichen. Insgesamt gaben die Versuchspersonen bei *kogkons* mehr korrekte Antworten, wobei die Differenz von nachher und vorher erzielten Punkten bei *bekog* höher liegt. Betrachtet man speziell die Aufgaben zum Faktenwissen (Wiedergabe von Formeln) und zum Problembewusstsein (Welche Punkte sind kritisch?), lässt sich feststellen, dass für das Faktenwissen bei beiden Systemen nahezu gleiche Werte ermittelt wurden. Bei den ‚Problemfragen‘ lässt sich in Hinblick auf den Punktezuwachs ähnliches feststellen. Betrachtet man jedoch die Teilantworten im einzelnen, beobachtet man, dass ein Themenkomplex, der bei *kogkons* im Gegensatz zu *bekog* besonders viel Initiative der Benutzer erforderte, auch mit Abstand am häufigsten in den Antworten genannt wurde. Offenbar bewährt sich die in *kogkons* angewendete Strategie, ein Problem von den Versuchspersonen aufarbeiten zu lassen in diesem Fall besonders gut. Bei *bekog* stand dagegen der am häufigsten genannte Punkt im Zusammenhang mit einer Aufgabe, deren Lösung als besonders schwierig erwies, weil der Sachverhalt zuvor offensichtlich zu knapp dargestellt worden war und die Benutzer dadurch genau auf das wesentliche Problem stießen. Es wäre sicherlich empfehlenswert, diese Strategie verstärkt anzuwenden, wobei gleichzeitig die Hilfe bei der Aufgabenlösung auszubauen ist (s.o.), um Frustration zu vermeiden.

System	insgesamt		Faktenwissen		Problembewusstsein	
	erzielte Punkte	Punktezuwachs	erzielte Punkte	Punktezuwachs	erzielte Punkte	Punktezuwachs
Bekog	6.14	4.57	1.42	1.42	0.86	0.86
Kogkons	7.2	3.7	1.8	1.3	1.5	0.9

Tab. 1: Übersicht über ausgewählte Ergebnisse zum Lernerfolg (maximale Punktzahl: 13)

5.2 Ergebnisse der Evaluation anhand des Kriterienkataloges

Die Evaluation von US-amerikanischen und deutschen Lernsystemen, bei der acht Programme (vier amerikanische und vier deutsche) anhand des Kriterienkatalogs bewertet wurden, ergab eine Vielzahl von Unterschieden in den Bereichen Design, Interaktion und Navigation, Inhaltspräsentation und Didaktik, wobei diese u.U. auch durch das jeweilige Lehrstoffgebiet bedingt sind. Zu den Gegenstandsbereichen gehörten vor allem Themen aus der Informationstechnologie (Einführung in den Umgang mit dem Internet Explorer, Einführung in die Datenverarbeitung, Oracle Programmierung, Hypermedia),

Grundlagen der Elektrotechnik sowie Themen aus den Bereichen

US-amerikanische Programme	Deutsche Programme
<i>Design</i>	
<ul style="list-style-type: none"> • Leerbereiche: ca. 60% der Bildschirmseite frei, kurze Textzeilen • Räumliche Organisation des Bildschirmbereichs wenig komplex • Kurze Textabschnitte, häufige Aufzählungen • Intensiver Einsatz von Multimedia • Grafiken und Animationen zur Illustration und Unterhaltung eingesetzt • Intensiver Einsatz von stark kontrastierenden Farben, Hervorhebung von Textteilen. 	<ul style="list-style-type: none"> • Eine Bildschirmseite nahezu vollständig mit Inhalten gefüllt, lange Textzeilen • Räumliche Organisation des Bildschirmbereichs komplex aber übersichtlich • Lange Fließtextabschnitte • Text ist wichtigstes Präsentationsmedium • Grafiken oder Simulationen als Lernhilfen eingesetzt • Gemäßigter Einsatz von Farben, Kontrast hauptsächlich zwischen einzelnen Bildschirmbereichen (Navigationsleiste, Inhaltsverzeichnis, Inhalt).
<i>Interaktion und Navigation</i>	
<ul style="list-style-type: none"> • Größere Auswahl an Interaktionsmöglichkeiten bei Übungen (Simulationen, Drag & Drop-Aufgaben, multimediale Elemente als Vorlagen für Fallstudien) • Eher beschränkte Navigationsfreiheit (Guided Tour, statisches Inhaltsverzeichnis, wenige Links zum Glossar) 	<ul style="list-style-type: none"> • Teilw. hohe Interaktivität durch intensiven Einsatz von Simulationen, teilw. gemäßigte Interaktivität bei Übungen (Multiple Choice, freie Eingabe), insg. häufiger Einsatz von Pop-Up-Fenstern im Inhaltsteil. • Größere Auswahl an Navigationsmöglichkeiten (Guided Tour, teilw. dynamisches Inhaltsverzeichnis, Browsing, Suchfunktion, Positionsanzeige, anspruchsvolles Hilfesystem zur Programmbedienung)

Tab. 2a: Merkmale US-amerikanischer und deutscher Lernprogramme im Vergleich

Pädagogik und Soziales (Problemlösungstechniken, Verhandlungsschulung, Soft Skills-Training). Einige der Merkmale (hier hervorgehoben) können auf die Einflüsse des in Abschnitt 2.2. beschriebenen sachsenischen und teutonischen wissenschaftlichen Stils (Galtung 1981) sowie der von Clyne (1991, 1994) definierten Unterschiede zwischen deutschen und englischen Diskursstrukturen zurückgeführt werden. Die folgende Tabelle gibt einen Überblick über die festgestellten Programmmerkmale, wobei dieser Vergleich ein Ergebnis darstellt, das mit Fortführung der Evaluation aktualisiert wird.

US-amerikanische Programme	Deutsche Programme
Inhalt	
<ul style="list-style-type: none"> • Detaillierte Angabe von Lernzielen und intensiver Einsatz von Advance Organizern vor jeder Lektion. • Fakten, Beispiele, Fallstudien, Angabe von Schrittfolgen und Problemlösungsstrategien, Checklisten (Do's & Don'ts) • Persönlicher Dialog mit dem Lernenden • Leicht verständliche Sprache, • kurze Sätze • Knappe Informationsdarstellung • Lineare bzw. flach hierarchische Informationsstruktur • Inhaltliche Unabhängigkeit einzelner Lektionen 	<ul style="list-style-type: none"> • Detaillierte Einführung in die Thematik, Inhaltsübersicht und allgemeine Lernzielangabe zu Beginn. • Theorie, Fakten, Beispiele, zu Anfang historischer Überblick. • Überwiegend unpersönliche Inhaltspräsentation • Komplexe Sprache, verschachtelte Sätze • Detaillierte Informationsdarstellung • Tief hierarchische Informationsstruktur • Inhalte einzelner Lektionen bauen aufeinander auf
Didaktik	
<ul style="list-style-type: none"> • Regelmäßiger Wechsel zwischen kurzen Lernstoffeinheiten und Übungen innerhalb einer Lektion • Aufgabentypen: Ja/ Nein – Fragen, Multiple Choice, Fallstudien • Ausführliche Vorgaben zur Aufgabenbearbeitung • Feedback: richtig/ falsch, Zusatzerläuterungen • Betonung von Werten wie Leistung, materielle Vorteile, praktische Anwendung, häufige Verwendung von Begriffen wie <i>nützlich, schnell, leicht, zeitsparend, effektiv</i> (extrinsische Motivation) 	<ul style="list-style-type: none"> • Übungen werden am Ende einer Lektion oder im separaten Übungsteil angeboten • Aufgabentypen: Multiple Choice, freie Eingabe (komplexe Verständnisfragen), Experimentiermöglichkeit durch den Einsatz von Simulationen • Knappe Vorgaben zur Aufgabenbearbeitung • Feedback: richtig/ falsch, Wiederholung des Lehrstoffs • Betonung von Werten wie Wissenserwerb, Verständnis, Qualität der Inhalte, Spaß (intrinsische Motivation)

Tab. 2b: Merkmale US-amerikanischer und deutscher Lernprogramme im Vergleich

5.3 Ergebnisse der Befragung

Die Auswertung der Umfragedaten, die im Kontext der kulturvergleichenden Lernstilanalyse erhoben wurden, ergab Unterschiede in den Lernstilen von Studierenden mit unterschiedlichem kulturellem Hintergrund. Die im Folgenden dargestellten Lernstilprofile von Studierenden aus ausgewählten Kulturen weisen erste Tendenzen im Hinblick auf die vermuteten kulturbedingten Unterschiede in Lernverhalten auf. Befragt wurden insgesamt 78 Studenten der Universität Hildesheim, wobei sich die Anzahl der Befragten aus den hier ausgewählten Kulturen zwischen drei und 17 verteilt.

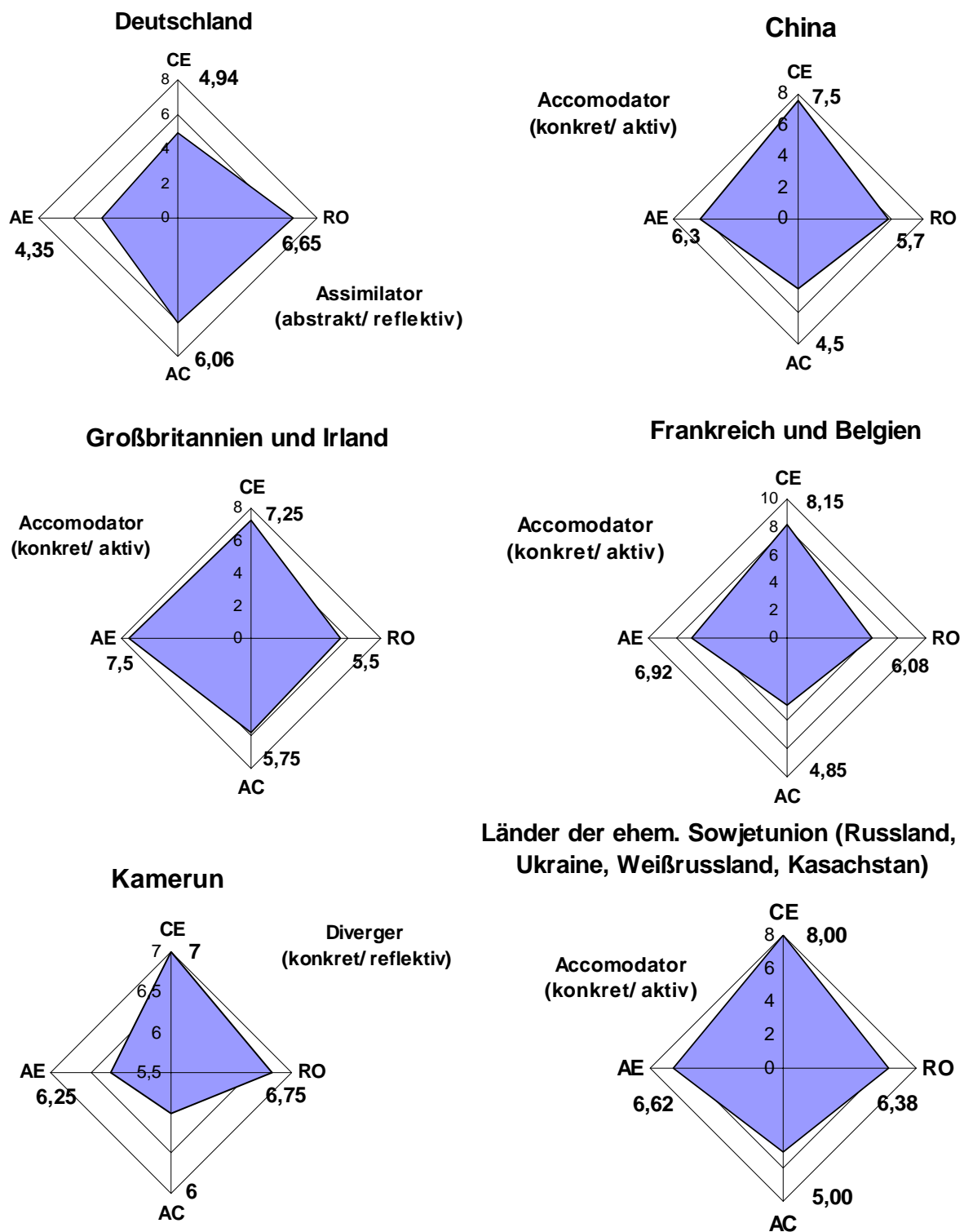


Abb. 1: Lernstilprofile von Studierenden aus ausgewählten Kulturen

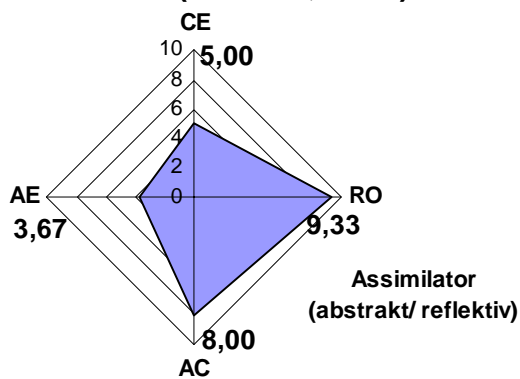
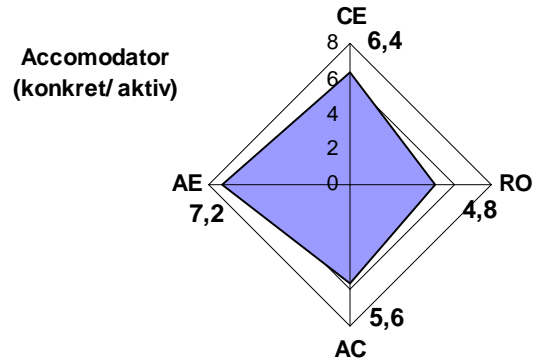
Südamerika (Bolivien, Peru)**Spanien**

Abb. 2: Lernstilprofile von Studierenden aus ausgewählten Kulturen (Fortsetzung)

In dieser Studie sind wir von der These ausgegangen, dass der Lernstil des Einzelnen nicht nur auf individuelle Persönlichkeitsmerkmale, sondern auch auf kulturelle Einflüsse, und insbesondere auf die des jeweiligen wissenschaftlichen Stils zurückgeführt werden kann. Nach der Auswertung des uns bisher vorliegenden Datenmaterials kann Galtungs Zuordnung der Kulturen zu den von ihm definierten akademischen Stilen nicht in jedem Fall auf die Lernstile übertragen werden. Das Profil der Studierenden aus den südamerikanischen Ländern, die Galtung als zum gallischen Stil zugehörig zählt, ähnelt beispielsweise dem Lernstilprofil deutscher Studierender. Die Studierenden aus den Ländern der ehemaligen Sowjetunion weisen einen ähnlichen Lernstil wie die chinesischen Studenten auf, obwohl sich laut Galtung diese Länder durch den teutonischen Stil auszeichnen. Diese Ergebnisse lassen vermuten, dass die Einflüsse des jeweiligen kulturspezifischen Bildungssystems in Bezug auf das Lernverhalten des Einzelnen von den allgemeinen gesellschaftlichen Werten und Überzeugungen innerhalb einer Kultur überdeckt werden können.

6 Fazit

Eine benutzergerechte Gestaltung von Lernsystemen äußert sich in der Berücksichtigung softwareergonomischer Faktoren. Eine Steigerung der Effektivität beim Lernen lässt sich zusätzlich durch die bewusste Einbeziehung lerntheoretischer Grundlagen erzielen, die ebenfalls auf die Benutzerbedürfnisse, aber auch auf Lerninhalt und -ziel abzustimmen ist. Das Projekt SELIM zielt auf die Verbindung beider Aspekte ab. In einem ersten Schritt sollen für ein eingegrenztes Lernziel (= Wiederholen und Einüben von Inhalten aus der Vorlesung) Gestaltungsmaßnahmen formuliert werden, die die Entwicklung von Benutzeroberflächen in diesem Bereich unterstützen. Zu diesem Zweck

erfolgte - ausgerichtet auf dieses Lernziel - die Konzeption zweier Systemprototypen mit unterschiedlicher lerntheoretischer Ausrichtung für einen konkreten Inhalt. Durch die Kombination von Aspekten verschiedener lerntheoretischer Ansätze sollen Schwächen kompensiert werden, die sich bei der programmtechnischen Umsetzung dieser Theorien ergeben. Bei einem vergleichenden Benutzertest beider Systeme ergaben sich erste Hinweise auf wichtige Aspekte der Gestaltung von Benutzeroberflächen (bezogen auf die Lehrform ‚Übung‘), wie bspw. die Erfordernis von Personalisierungsmöglichkeiten hinsichtlich der optischen Gestaltung. Diese Erkenntnisse sind durch eine Übertragung auf weitere Inhalte zu überprüfen und evtl. zu ergänzen.

Zusätzlich wird bedacht, dass Lernsituationen in hohem Maße individualisiert sind, sodass Lernsysteme durch eine angemessene Aufbereitung des Lernstoffes auf unterschiedliche Fähigkeits- und Erfahrungsebenen reagieren und individuelle Navigationsunterstützung und Darstellung der Inhalte bieten sollten. Die Individualisierung des Lernens wird durch die Führung eines adaptiven Dialogs zwischen Benutzer und System erreicht. Voraussetzung ist dabei eine Lernermodellierungskomponente mit diagnostischen Funktionen, die den aktuellen Stand des Lernprozesses ermitteln kann. Handelt es sich bei dem Kreis der Systemnutzer um eine multikulturelle Gruppe, umfasst der Kontext der Lernsituation nicht nur das individuelle Vorwissen oder die Interessen der Benutzer, sondern auch deren kulturellen Hintergrund, der sich auf ihr Vorgehen während des Lernprozesses auswirkt. Hieraus ergibt sich die Notwendigkeit der Erweiterung der Adaptationsleistung von interaktiven Lernsystemen um den Aspekt der kulturellen Orientierung. Mit unserer Studie soll ein Beitrag geleistet werden zur Erforschung kultureller Voraussetzungen für das multimediale Lernen, die als Basis für die Gestaltung der Benutzermodellierungskomponente eines adaptiven Lernsystems dienen soll. Hierzu werden zunächst Merkmale ermittelt, die für Lernprogramme aus verschiedenen Kulturen charakteristisch sind, insbesondere in den Bereichen Design, Interaktions- und Navigationsmöglichkeiten, Inhaltspräsentation und Didaktik. Die kulturvergleichende Lernstilanalyse dient der Ergänzung dieser Ergebnisse.

7 Literatur

- Barmeyer, Ch. (2000). „Interkulturelles Management und Lernstile. Studierende und Führungskräfte in Frankreich, Deutschland und Quebec“. Frankfurt et al.: Campus Verlag.
- Baumgartner, P.; Payr, S. (1997). „Erfinden lernen“. In: Müller, K.H.; Stadler, F. (Hrsg.): Konstruktivismus und Kognitionswissenschaft. Kulturelle Wurzeln und Ergebnisse. Wien et al.: Springer. 89-106

- Beneke, J. (2001). "The 14 Dimensions of Culture. Orientation Matrix". Research Centre for Intercultural Communication, University of Hildesheim, Hildesheim.
- Blumenstengel, A. (1998). „Entwicklung hypermedialer Lernsysteme“. Berlin: Wissenschaftlicher Verlag.
- Brusilovsky, P. (1998). "Methods and Techniques of Adaptive Hypermedia." In: Brusilovsky, P.; Kobsa, A.; Vassileva, J. (Eds.). Adaptive Hypertext and Hypermedia. Boston et al.: Kluwer Academic Publishers.
- Caroli, F. (2000). „Internetgestützte Seminare. Innovation für eine Form der universitären Lehre.“ In: Scheuermann F. (Hrsg.). Campus 2000. Lernen in neuen Organisationsformen. Münster: Waxmann. 149-160.
- Clyne, M. (1991). "The Sociocultural Dimension: The Dilemma of the German-speaking Scholar." In: Schröder, H. (Hrsg.). Subject-oriented Texts. Languages for Special Purposes and Text Theory. Berlin/New York: de Gruyter.
- Clyne, M. (1994). "Inter-cultural communication at work. Cultural values in discourse." Cambridge: Cambridge University Press.
- Galtung, J. (1981). "Structure, culture and intellectual style: An essay comparing saxon, teutonic, gallic and nipponic approaches." In: Social Science Formation. London/Beverly Hills: SAGE. 817-856.
- Hall, E. T.; Hall, M. R. (1990). "Understanding Cultural Differences: Germans, French and Americans". Yarmouth: Intercultural Press.
- Hofstede, G. (1986). "Cultural Differences in Teaching and Learning." In: International Journal of Intercultural Relations, Vol.10. 301ff.
- Hofstede, G. (1993). „Interkulturelle Zusammenarbeit - Kulturen, Organisationen, Management“. Wiesbaden: Gabler.
- Jonassen, D.H.; Grabowski, B.L. (1993). "Handbook of Individual Differences, Learning, and Instruction". Hillsdale: Lawrence Erlbaum.
- Kerres, M. (1998). „Multimediale und telemediale Lernumgebungen: Konzeption und Entwicklung“. München et al.: Oldenbourg.
- Kobsa, A.; Wahlster, W. (1989). "User Models in Dialog Systems." in: Kobsa, A., Wahlster, W.(Eds.). User Models in Dialog Systems. Berlin et al.: Springer.
- Kobsa, A.; Koenemann, J.; Pohl, W. (2000). "Personalized Hypermedia Presentation Techniques for Improving Online Customer Relationships." GMD FIT, St. Augustin. <http://fit.gmd.de/~kobsa/papers/Kobsa-PHPT-draft.doc> [03.08.2002]
- Kolb, D. (1984). "Experiential Learning". New York: Prentice Hall.
- Kolb, D.; Osland, J.; Rubin, I. (1995). "Organizational Behavior. An experiential approach." Englewood Cliffs: Prentice Hall.
- Krause, J.; Womser-Hacker, Ch. (1997). „Vages Information Retrieval und graphische Benutzungsoberflächen. Beispiel Werkstoffinformation.“ Konstanz: Universitätsverlag.
- Reinmann-Rothmeier, G.; Mandl, H. (Hrsg.) (2001). „Virtuelle Seminare in Hochschule und Weiterbildung. Drei Beispiele aus der Praxis.“ Bern et al.: Verlag Hans Huber.
- Rüppel, H.; Pfleging, B. (2000) „ALICE: Adaptives Lernen – Interaktiv, Kooperativ, Explorativ. Mehr Anwendung, mehr Training, weniger träges Prüfungswissen.“ (<http://www.uni-koeln.de/phil-fak/paedsem/psych/alice/konzept.pdf> [03.08.2002])

Elisabeth Kamentz, Monika Schudnagis

Tulodziecki, G.; Hagemann, W.; Herzig, B.; Leufen, S.; Mütze, C. (1996). „Neue Medien in den Schulen: Projekte-Konzepte-Kompetenzen.“ Gütersloh: Bertelsmann Stiftung.



U-Know - Ubiquitäres Wissensmanagement an der Universität Regensburg

Stefan Berger, Franz Lehner

Universität Regensburg
Universitätsstraße 31
D-93053 Regensburg
Tel.: 0941-943-3207
Fax: 0941-943-3211
stefan.berger@wiwi.uni-regensburg.de

Zusammenfassung

Die hohe Mobilität in der Informationsgesellschaft bedeutet auch für das Wissensmanagement neue Herausforderungen – etwa wenn Mitarbeiter auch unterwegs auf aktuelle Informationen ihres Unternehmens angewiesen sind. Diesem Umstand wird am Lehrstuhl für Wirtschaftsinformatik III unter Leitung von Prof. Dr. F. Lehner mit der Entwicklung eines mobilen Wissensmanagementsystems Rechnung getragen. Für die Wirtschaftswissenschaftliche Fakultät wird der Prototyp „U-Know“ („Ubiquitous Knowledge Management“) konzipiert, der die Nutzung von typischen Wissensmanagementfunktionen auf mobilen Geräten wie PDAs und Handys ermöglicht. Mitarbeiter der Fakultät können so ortsunabhängig auf benötigte Informationen zugreifen. Auf diese Weise wird der Wissensaustausch in der Fakultät gefördert und die Arbeitsabläufe in Forschung und Lehre gewinnen an Effizienz.

Abstract

This paper describes the concept of a mobile Information System which will be implemented at the Faculty of Business, Economics and Information Systems at the University of Regensburg. The prototype „U-Know“ („Ubiquitous Knowledge Management“) will enable typical functions of Knowledge Management Systems for mobile devices such as Mobile Phones and Personal Digital Assistants. In this way faculty staff can access information they need to fulfil their daily tasks while travelling or being on the move at the university campus.



1 Einführung

Steigender Wettbewerb und Internationalisierung sind Herausforderungen, die nicht nur Unternehmen, sondern in zunehmendem Maße auch die universitären Bildungseinrichtungen in Deutschland betreffen. Die Monopolkommission fordert beispielsweise in ihrem Sondergutachten zur Situation an deutschen Hochschulen mehr Wettbewerb als neues Leitbild für die Hochschulpolitik. Auch international nimmt die Konkurrenz stetig zu. Während deutsche Universitäten bis zum Ersten Weltkrieg weltweit eine Spitzenstellung innehaten, befinden sich die Hochschulen hierzulande in einem heute von vielen Seiten beklagten Zustand [Mono02]. Bildungseinrichtungen im angelsächsischen Raum sind für begabte Wissenschaftler und Studenten aus Deutschland und aller Welt häufig attraktiver.

Für Hochschulen gilt es, mit entsprechenden Strategien auf diese Entwicklungen zu reagieren, um den nationalen und internationalen Anschluss nicht zu verpassen. Dem zielgerichteten Umgang mit dem in der Administration vorhandenen Wissen über Prozesse, Verordnungen, Zuständigkeiten usw. kommt dabei eine zentrale Rolle zu. Die Nutzung dieses Metawissens kann die Transparenz und Effizienz in den Arbeitsabläufen und damit die Qualität von Forschung und Lehre verbessern. Zu diesem Zweck muss es allerdings systematisch erfasst und für die Mitarbeiter bereitgestellt werden. Wissensmanagement ist daher auch für Universitäten ein aktuelles Thema. Dieser Aufsatz präsentiert den Entwurf eines mobilen Wissensmanagementsystems für die Wirtschaftswissenschaftliche Fakultät der Universität Regensburg, das der Benutzerzielgruppe (Professoren, Mitarbeiter, Verwaltungsangestellte) über mobile Endgeräte einen ortsunabhängigen Zugriff auf benötigte Informationen ermöglichen soll. Zunächst wird auf die Frage nach Wissensmanagement und auf Wissensmanagementsysteme in der öffentlichen Verwaltung eingegangen (Kapitel 2). In Kapitel 3 wird aufgezeigt, wie sich der Faktor Mobilität auf das Wissensmanagement auswirkt und warum er beim Entwurf eines entsprechenden Wissensmanagementsystems für die Universität beachtet werden sollte. Kapitel 4 stellt schließlich die Konzeption des Prototypen U-Know (Ubiquitous bzw. Universitäres **Knowledge** Management) vor. Es wird aufgezeigt, welche Funktionen realisiert werden sollen und wie sich der Prototyp in bestehende Systeme und Anwendungen einfügt. Kapitel 5 fasst die wichtigsten Punkte noch einmal zusammen.

2 Wissensmanagement an der Universität

Zum Wissensmanagement existiert mittlerweile eine umfangreiche Fülle an Definitionen, Theorien, Konzepten und Modellen - und das obwohl das Thema vergleichsweise jung ist. Für einen Überblick über die verschiedenen Wissensmanagementansätze sei auf [Lehn00] verwiesen (dies gilt ebenso für die oft diskutierten Definitionen der Begriffe „Daten“, „Informationen“ und „Wissen“). An dieser Stelle soll zunächst der Frage nachgegangen werden, welche Ziele mit Wissensmanagement in der öffentlichen Verwaltung verbunden sein können und wie Wissensmanagement-Aktivitäten technisch unterstützt werden können.

2.1 Effizienz als Ziel des Wissensmanagements

Für Unternehmen im Informationszeitalter ist Wissen die Basis ihrer Wettbewerbsfähigkeit, da es die Innovationsgeschwindigkeit, die Effizienz von Prozessen, die Qualität von Produkten usw. bestimmt [BaÖs99]. Vor diesem Hintergrund versuchen zahlreiche Unternehmen, vorhandenes Wissen durch Wissensmanagementaktivitäten zu nutzen und neues zu entwickeln. Die Aktivitäten zielen beispielsweise darauf ab, unternehmensinterne Informationsflüsse zu verbessern, wichtige Ressourcen optimal zu nutzen, die Effizienz in F&E-Projekten zu erhöhen und die „Time-to-Market“ zu verkürzen. Wissen bedeutet Wettbewerbsvorteile [Lehn00].

Die Aufgaben einer (Uni-) Fakultät umfassen die Durchführung des Unterrichts, die Heranbildung des wissenschaftlichen und künstlerischen Nachwuchses, die Verantwortung für eine wirksame Studienberatung sowie die Sorge für die wissenschaftliche Forschung und die Anwendung hochschuldidaktischer Erkenntnisse [Baye98]. Da sie sich als Organ der öffentlichen Hand aufgrund einiger Besonderheiten von Unternehmen der Privatwirtschaft unterscheidet, sind die oben genannten Ziele auch nicht unbedingt auf sie übertragbar. Eine Universität verfügt über Zwangsmittel und kann einseitig handeln (z. B. bei der Exmatrikulation von Studenten). Sie sind deshalb in der Regel nicht auf Austauschbeziehungen mit ihren Adressaten sowie auf Gewinnerzielung angewiesen. Angesichts begrenzter Haushaltsmittel und knapper Ressourcen müssen jedoch auch Behörden versuchen, Arbeitsabläufe möglichst effizient zu gestalten. Die Verfolgung von Effizienzzielen gilt umso mehr für Universitäten, die ja zunehmend im nationalen und internationalen Wettbewerb stehen (s. oben). Effizienzgewinne ergeben sich beispielsweise, wenn

- sich Durchlaufzeiten bei Verwaltungsvorgängen verkürzen,
- Zeit- und Kosteneinsparungen stattfinden
- und eine verstärkte Kundenorientierung (z. B. durch eine schnellere Auskunft) erreicht wird.

Kennzeichnend für die Organisationsstruktur der Universität ist, dass das vorhandene Wissen an unterschiedlichen Stellen verteilt ist, da die Forschung und Lehre weitgehend dezentral von den Lehrstühlen und die Verwaltung von der Fakultätsleitung (Dekanat) verantwortet wird. Jedoch besteht zur Durchführung der oben genannten Aufgaben umfangreicher Koordinations- und Kommunikationsbedarf der beteiligten Stellen, wie er in Abbildung 1 angedeutet ist.

Die Arbeitszeit, die für administrative Tätigkeiten aufgewendet wird, ist dabei beträchtlich. Gerade, weil die genauen Abläufe aufgrund der relativ großen Autonomie der Beteiligten unklar sind, stellt sich z. B. häufig die Frage, wer die richtigen Ansprechpartner in Verwaltungsangelegenheiten sind. Als Folge können Verzögerungen und damit Ineffizienzen auftreten.

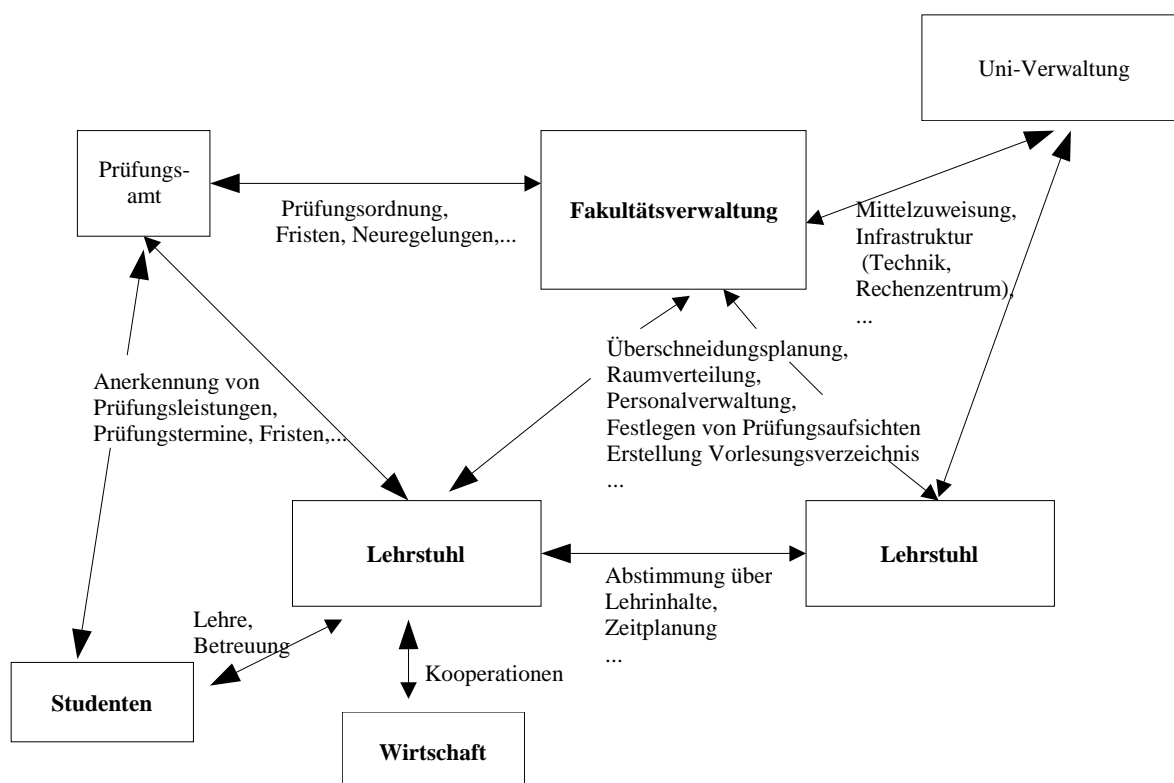


Abbildung 1: Koordinationsaufwand an der Fakultät

Vor diesem Hintergrund hat Wissensmanagement eine besondere Relevanz, da angesichts der These, dass nur 20 bis 30 Prozent des eigentlich verfügba-

ren organisatorischen Wissens wirklich genutzt werden, noch große Produktivitätspotenziale durch zielgerichtetes Wissensmanagement vermutet werden [Schü96].

Die Erreichung der Ziele kann durch Wissensmanagementsysteme unterstützt werden, welche die Erfassung und den Zugriff auf benötigtes Wissen erleichtern. Wissensmanagement umfasst als Managementfunktion [Maie02] natürlich weitaus mehr als technische Aspekte. Es ist jedoch zu beobachten, dass Wissensmanagementsysteme häufig den ersten Schritt zu einem zielgerichteten Wissensmanagement darstellen. Auch hier gibt es mittlerweile ein großes Angebot von Systemen mit unterschiedlicher Ausrichtung und umfangreichen Funktionen. Der nächste Abschnitt versucht, an Beispielen den Stand der IT-Unterstützung des Wissensmanagements in Universitäten aufzuzeigen.

2.2 Wissensmanagement an Universitäten

Auch an Universitäten existieren erste Ansätze zur Implementierung eines Wissensmanagements. Häufig werden die Aktivitäten softwaretechnisch unterstützt bzw. umgesetzt. Beispiele hierfür sind:

- Groupware-Anwendungen und Dokumenten-Management-Systeme (z.B. Lotus Notes / Domino an der **Universität St. Gallen**, wo neben der Mail- und Gruppenkalenderfunktion, Dokumentenverwaltung auch elektronische Anschlagbretter, ein Mitarbeiterhandbuch und Vorlesungs- und Veranstaltungskalender im Einsatz sind),
- Data Warehouse-Lösungen (z.B. das an der **Universität Bamberg** entwickelte Informations- und Berichtssystem CEUS-HB, das u. a. die Forschungsberichterstattung, die Mittelverwaltung und die Personalwirtschaft unterstützt),
- Intranet-Lösungen (z.B. das Intranet an der Wirtschaftswissenschaftlichen Fakultät der Uni Regensburg, wo u.a. aktuelle Ankündigungen und häufig benötigte Formulare bereitgestellt werden).

Neben diesen Anwendungen wäre auch der Einsatz einer eigenen Klasse von Werkzeugen zur Unterstützung des Wissensmanagements denkbar, die unter dem Begriff Wissensmanagementsysteme (WMS) zusammengefasst werden. Sie kombinieren und integrieren beispielsweise die Funktionalitäten der oben genannten Lösungen [Maie02]. Durch derartige Systeme sollen die Wissensmanagementziele schneller oder besser erreicht werden [Klos01]. Bisher ist dem Autor allerdings noch kein Beispiel bekannt, wo Wissensmanagement an

einer Universität systematisch durch den Einsatz eines WMS unterstützt wird. Damit sind die Wissensmanagementaktivitäten an Universitäten eher als im Anfangsstadium befindlich zu bezeichnen.

Gemeinsam ist den meisten Systemen, die im Wissensmanagement eingesetzt werden, dass sie überwiegend als Desktop-Systeme konzipiert und damit auf den stationären Gebrauch ausgerichtet sind. Der Faktor Mobilität wurde bisher in der Entwicklung von Informationssystemen nicht berücksichtigt [BeBl96;Fagr99]. Mitarbeiter, die im Rahmen ihrer Aufgaben mobil sind, können daher nur beschränkt auf die vorhandenen Informationsressourcen zugreifen. Daraus entstehen den mobilen Mitarbeitern Informationsnachteile, wie im Kapitel 3 gezeigt wird.

3 Mobilität als Herausforderung für das Wissensmanagement an der Universität

Die hohe Mobilität ist kennzeichnend für die Gesellschaft im Informationszeitalter [Lehn01], wo geografische Entfernungen nur noch eine untergeordnete Rolle spielen: „In the modern world remoteness is not about distance, but about information.“ [DiBe96] Mobilität wirkt sich unmittelbar auf das Wissensmanagement aus, denn viele Mitarbeiter sind im Rahmen ihrer täglichen Aufgaben mobil. Sie benötigen auch unterwegs Zugriff auf die Wissensbasis ihrer Organisation. Dies trifft ebenso für die Angehörigen einer Universität zu: Professoren und wissenschaftliche Mitarbeiter reisen beispielsweise zu Konferenzen oder sind ebenso wie Verwaltungsangestellte auf dem Universitätsgelände unterwegs. Der Wissensbegriff erhält zwar durch den Zusatz „mobil“ keine neue (Be)Deutung, aber Mobilität stellt eine neue Gestaltungsdimension im Kontext von Informationssystemen dar. Die Bezeichnung als „mobiles Wissen“ kann in verschiedenen Zusammenhängen gesehen werden:

1. Mobiler Zugriff bzw. mobile Zugriffsmöglichkeit auf Wissen und Informationen
2. Mobilität des Wissens durch Mobilität der Wissensträger
3. Mobilität des Wissens durch Mobilität des „Entstehungsortes“ (z. B. bei Verkehrsinformationen)
4. Der Prozess, in den das Wissen eingebunden ist, findet an unterschiedlichen Orten statt (z. B. im Beratungsgeschäft, aber auch bei Wartungs- und Serviceleistungen)

Gemeinsam ist allen mobil Arbeitenden, dass sie Transaktionen oder Teile von Transaktionen in Bewegung oder an wechselnden Aufenthaltsorten durchführen, womit eine raum-zeitliche Entkopplung von stationären Kommunikationspartnern verbunden ist [Schu99]. Für eine Organisation und ihre Mitarbeiter ergeben sich dadurch eine Reihe von Problemen, wie etwa [BeBl96;Froh94;Perr01]:

- Unterbrechungen im Informationsfluss, woraus Zeitverzögerungen in Arbeitsabläufen resultieren,
- eingeschränkte „Zugriffsmöglichkeit“ auf das spezifische Wissen der mobilen Mitarbeiter durch ihre schlechte Erreichbarkeit (z. B. ist der Kontakt über Email eingeschränkt), wenn sie sich nicht an ihrem Arbeitsplatz befinden,
- mangelnde Kenntnis und Mitführbarkeit aller vor Ort benötigten Informationsressourcen,
- eingeschränkter informeller Kontakt zu Kollegen, der von wesentlicher Bedeutung für Koordination, Teamgeist und Problemlösungskompetenz ist (beispielsweise ist es nicht ohne weiteres möglich, in einem zwanglosen Gespräch Erfahrungen mit den anwesenden Kollegen auszutauschen oder bei unbekannten Lösungen nach Experten zu suchen).

Diese Probleme können zu Effizienzverlusten in den Arbeitsabläufen führen. Die Mobilität ihrer Arbeitnehmer stellt somit auch für Universitäten eine Herausforderung im Wissensmanagement dar. Sie besteht darin, Mobilität als einen Prozessparameter bei der Gestaltung der Arbeitsabläufe in Betracht zu ziehen [Bind01]. Durch den Einsatz mobiler Informations- und Kommunikationstechnologien kann der Faktor Mobilität bei der Prozessgestaltung berücksichtigt werden. An der Universität Regensburg wird daher an einem Prototyp für mobiles Wissensmanagement gearbeitet, der speziell auf die Nutzung durch mobile Endgeräte abgestimmt ist.

4 Mobiles Wissensmanagement mit „U-Know“

Am Lehrstuhl für Wirtschaftsinformatik III entsteht unter Leitung von Prof. Dr. F. Lehner der Prototyp U-Know, mit dem ab September 2002 erste Wissensmanagement-Funktionen mobil nutzbar sein sollen. Die Systementwicklung erfolgt evolutionär. Bei diesem Vorgehen, auch als Versioning bezeichnet, wird das Produkt als Folge von Versionen betrachtet, so dass jede Version gewissermaßen der Prototyp für die Folgende ist [LeH+95]. Dies stellt si-

cher, dass U-Know relativ zügig genutzt werden kann und aus der Nutzung resultierende neue Anforderungen sukzessive implementiert werden können.

Durch die Arbeit werden zwei Forschungsschwerpunkte des Lehrstuhls, nämlich Mobile Business und Wissensmanagement, verbunden. U-Know wird zunächst für die Wirtschaftswissenschaftliche Fakultät der Universität Regensburg konzipiert. Die Sicht auf U-Know wird in die Ebenen „Wissensbasis“, „Systemfunktionen“ und „Benutzer & Zugriff“ unterteilt, die aufeinander aufbauen (s. Abbildung 2). Im Folgenden werden diese Ebenen einer näheren Betrachtung unterzogen.

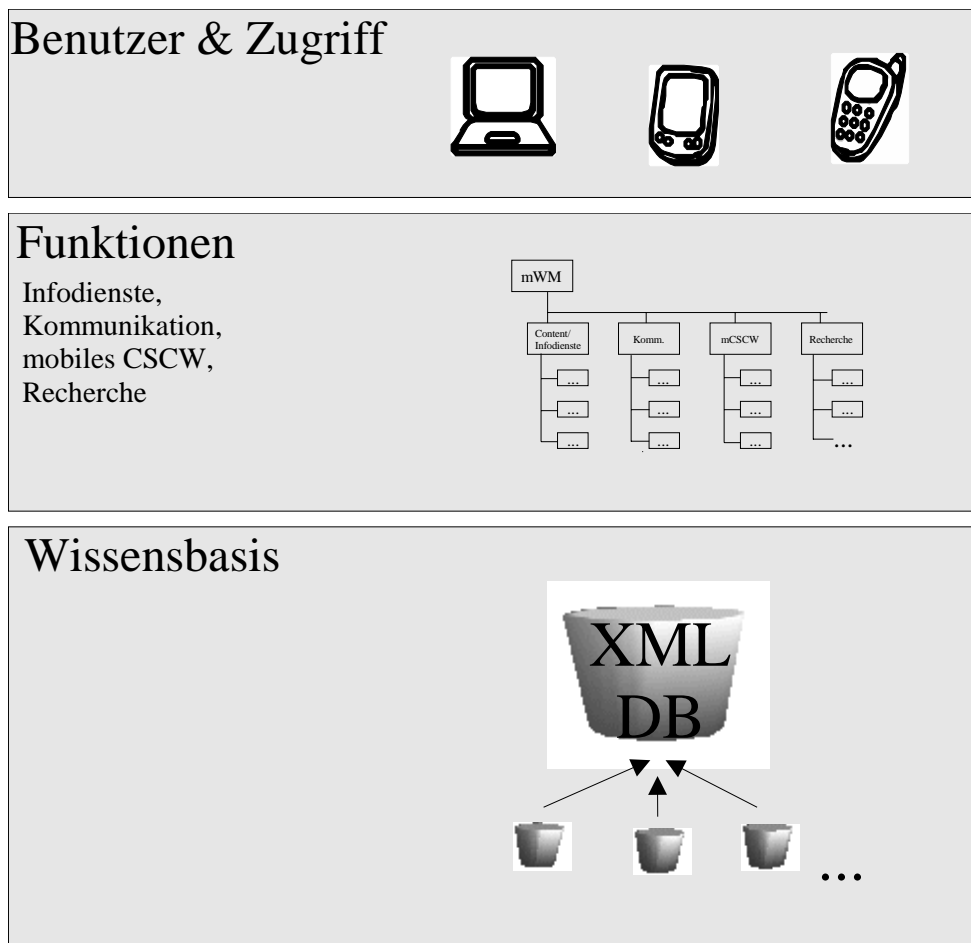


Abbildung 2: Ebenen von "U-Know"

4.1 Wissensbasis

In einem ersten Schritt gilt es, die benötigten Daten, in einer einheitlichen Wissensbasis zusammenzuführen. Bisher sind die Daten und Informationen, die in U-Know enthalten sein sollen, in Datenbanken verschiedener Fakultäts-Server verfügbar. Dabei handelt es sich gewissermaßen um unverbundene

Wissensinseln, die Teil des Wissensraums¹ der Fakultät sind (vgl. Abbildung 3). Werden spezielle Informationen benötigt, muss man auf der Suche nach dokumentierten Wissen oder Ansprechpartnern die verschiedenen Server einzeln durchsuchen. Die Daten werden daher aus den verschiedenen relationalen Datenbanken extrahiert. Hierzu ist es erforderlich, den Prototypen in die gewachsene Serverlandschaft der Fakultät, die in Abbildung 4 dargestellt wird, zu integrieren.

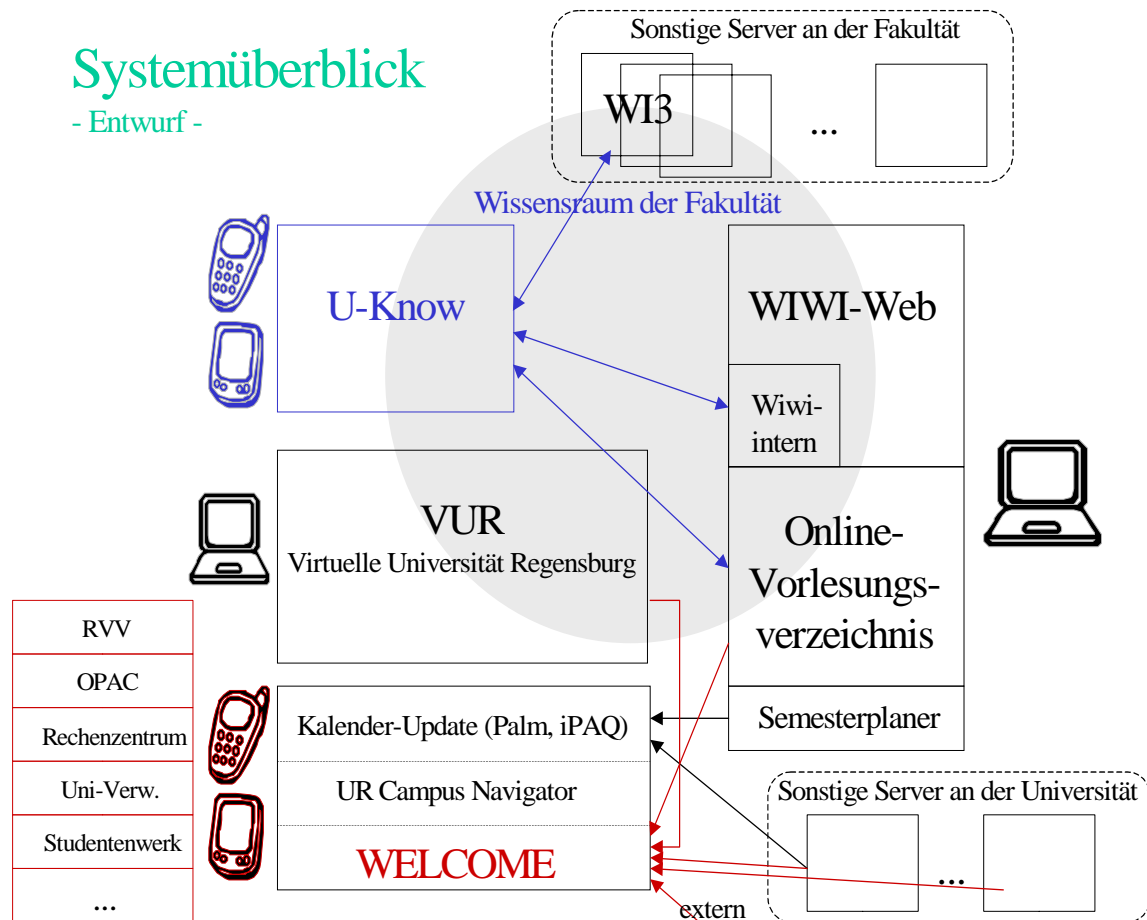


Abbildung 3: Systemüberblick: Server an der Wiwi-Fakultät

- WIWI-Web

Mit „WIWI-Web“ wird der Webserver der Fakultät bezeichnet. Hier sind folgende für U-Know relevante Informationen vorhanden:

- aktuelle Mitteilungen,
- die Lehrstühle mit Inhabern und vertretenem Fachgebiet,
- Kontaktadressen (Anschriften und Telefonnummern der Verwaltungsorgane),

¹ Zum Begriff des „Wissensraums“ vgl. [Lehn00], S. 141.

- Professoren (mit Links zu den Lehrstühlen),
- Personal der Fakultät (Professoren, wissenschaftliche Mitarbeiter, Verwaltungs-angestellte),
- Verwaltung (mit dem Verzeichnis der Verwaltungsorgane, Stellenausschreibungen und häufig benötigten Formularen wie Reise- oder Urlaubsanträge usw.).
- Online-Vorlesungsverzeichnis
In Zukunft soll im Online-Vorlesungsverzeichnis das gesamte Vorlesungsangebot der Fakultät verfügbar sein. Diese Informationen werden beispielsweise benötigt um festzustellen, wann gewünschte Ansprechpartner verfügbar sind. Daneben sind die Fakultätsangehörigen mit Telefon- und Raumnummer erfasst.
- sonstige Server der Fakultät
Neben dem Web-Auftritt der Fakultät sind auch alle Lehrstühle mit einem eigenen Web-Auftritt vertreten. An für U-Know relevanten Informationen finden sich hier weitere Angaben zu den Forschungsschwerpunkten, Projekten und Mitarbeitern der einzelnen Lehrstühle.

Die Informationen werden in einer XML-Datenbank („Extensible Markup Language“) als XML-Dokumente gespeichert. Beim XML-Standard handelt es sich um eine Metasprache mit der die Dokumentinhalte in einer vorher definierten Struktur abgelegt und die verschiedenen Elemente des Inhalts durch sogenannte Tags gekennzeichnet (daher Markup Language = Auszeichnungssprache) werden. Die Struktur wird vom Benutzer individuell entworfen und separat in einer sogenannten DTD (Document Type Definition) gespeichert. Ein Vorteil bei der Verwendung von XML besteht in der Trennung von Layout und Inhalt. Damit ist es möglich die Repräsentation der Dokumente auf das jeweilige Endgerät, mit dem der Zugriff erfolgt, abzustimmen. Beispielsweise können dadurch aktuelle Mitteilungen auf einem PDA in voller Länge und auf dem kleinen Display eines Handys dagegen lediglich als Überschrift mit Datum dargestellt werden.

4.2 Funktionen von U-Know

Das Ziel von U-Know ist, typische Funktionalitäten von WMS durch mobile Technologien zu unterstützen. Die genaue Ausgestaltung von WMS ist recht unterschiedlich, so dass man nicht von einem typischen WMS oder „Referenz“-WMS sprechen kann. Allerdings existieren bestimmte funktionale Anforderungen, die immer wieder genannt werden [Lehn00; Klos01]. Dabei handelt es sich u. a. um Funktionen zur

- **Wissensrepräsentation**
Dieser Funktionsbereich fasst Funktionen zusammen, welche die Darstellung der Wissenselemente betreffen. Wissenslandkarten und Gelbe Seiten aber auch To Do Listen bieten sich z. B. an, um die Inhalte zu strukturieren und Zusammenhänge zwischen verschiedenen Wissensgebieten zu verdeutlichen. Damit erleichtern sie wiederum die Navigation in einem WMS.
- **Wissenskommunikation und –kooperation**
„Ohne Kommunikation kein common knowledge [Lehn00]“. Die persönliche Kommunikation und Kooperation der Mitarbeiter ist ein zentraler Prozess für die Schaffung von Wissenselementen. Daher sollten WMS die Zusammenarbeit und Kommunikation in Teams unterstützen. Hier handelt es sich um einen Bereich, der aus Groupware-Anwendungen schon bekannt ist, aber in den Wissenskontext integriert werden kann [Klos01]. Typische Funktionen sind z. B. Diskussionsforen, Workflow Management, Terminfinder, Gruppenkalender usw.
- **Wissenssuche**
Um an die gewünschten Informationen zu kommen, muss die Navigation durch den Datenbestand im System ebenso wie die Schlagwortsuche nach Begriffen möglich sein. Über Filter ist die Suche auf bestimmte Themengebiete, Zeiträume, Dateiformate usw. beschränkbar. Über Hyperlinks werden die einzelnen Wissenselemente vernetzt. Zusätzlich können intelligente Agenten verwendet werden, die dem Benutzer Informationen zu relevanten Themengebieten oder Änderungen zustellen („Push-Technologie“).

Diese Funktionen sollen im Rahmen von U-Know für mobile Endgeräte realisiert werden. Auf dem Markt befindliche WMS beschränken sich bisher zumeist auf die Möglichkeit zum Browsen der Verzeichnisstruktur mittels eines WAP-fähigen Mobiltelefons und SMS-Benachrichtigungsdienste. Aus Funktionssicht umfasst der Prototyp Hauptfunktionen, die in weitere Teilfunktionen aufgespaltet werden (s. Abbildung 4).

Von besonderem Interesse für Benutzer sind die speziell für mobile Endgeräte gedachten Funktionen „mobiles Wissensmanagement“, welche an dieser Stelle vorgestellt werden sollen. Damit werden die eigentlichen Funktionen eines WMS realisiert (auf Funktionen, die beispielsweise die Administration von U-Know betreffen, wird hier nicht eingegangen).

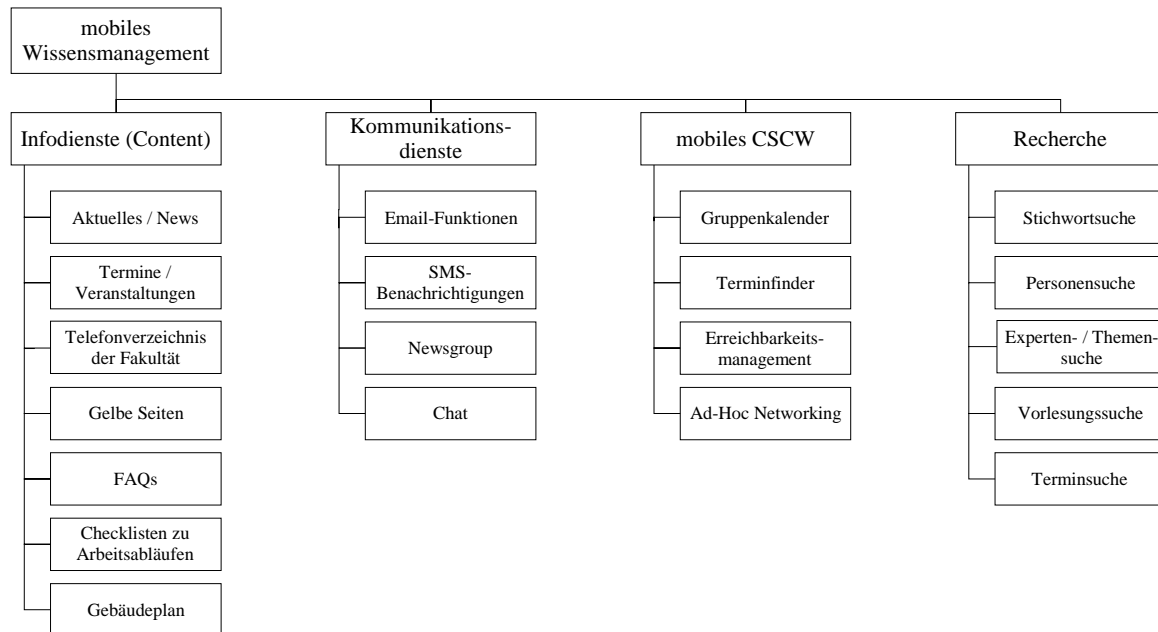


Abbildung 4: Mobile WM-Funktionen von "U-Know"

- **Infodienste**

Über verschiedene „Infodienste“ kann auf die Wissensobjekte des Systems zugegriffen werden (Funktion Wissensrepräsentation). Die Struktur beruht auf dem Prinzip von Wissens(land)karten, die das in einer Organisation vorhandene Wissen textuell und / oder grafisch in gegliederter Form darstellen [Lehn00]. Der Benutzer findet Informationen zu Neuigkeiten, Terminen und den Organen der Fakultät. Einen wichtigen Teil von Wissenslandkarten bilden gewöhnlich die Gelben Seiten. Sie sollen einen direkten Kontakt mit jenen Personen ermöglichen, die über benötigtes Wissen verfügen. Der Wissensaustausch in der Fakultät wird erleichtert und Doppelarbeiten dadurch unter Umständen vermieden. Eine FAQ-Liste kann dazu verwendet werden, weniger strukturiertes Wissen zu erfassen [Lehn01] (z.B. „Wer ist für den Email-Verteiler der Fakultät zuständig?“). Für standardisierte Arbeitsabläufe (z.B. Einstellung eines neuen Mitarbeiters) eignen sich Checklisten, die den aktuellen Bearbeitungsstand anzeigen. Ein Gebäudeplan erleichtert darüber hinaus z. B. für neue Mitarbeiter die Orientierung, wenn Ansprechpartner zu einem bestimmten Thema gesucht werden („Wohin muss ich, wenn ich eine Rechnerkennung benötige?“).

- **Kommunikationsdienste / mobiles CSCW**

„Kommunikationsdienste“, wie etwa der mobile Zugriff auf Emails oder SMS-Benachrichtigungen sowie „mobile CSCW-Funktionen“ sollen die persönliche Kommunikation erleichtern. Sie ist trotz technischer Unterstützung des Wissensmanagements unentbehrlich für den Wissensaustausch der Mitarbeiter. Über ein Erreichbarkeitsmanagement können Benutzer z.B. abrufen, wann und auf welchem Weg der gewünschte Ge-

sprachspartner am besten verfügbar ist. Ad-Hoc Networking zeichnet sich dadurch aus, dass die mobilen Geräte ein temporäres Netzwerk bilden, ohne auf eine bestehende stationäre Infrastruktur oder zentrale Administration angewiesen zu sein [Hela99]. Damit können beispielsweise elektronische Visitenkarten oder Präsentationsunterlagen ausgetauscht werden.

- **Recherchefunktionen**

Neben der Navigation durch die Wissensstruktur sollten die vorhandenen Wissensobjekte auch über verschiedene Recherchemöglichkeiten auffindbar sein. Damit kann nach Stichworten, Personen, Themen, Vorlesungen oder Terminen gesucht werden. Von besonderer Bedeutung ist in diesem Zusammenhang auch die Möglichkeit zur „Expertise Location“ [McDAc98], also die Suche nach Experten zu einem bestimmten Fachgebiet, falls unterwegs Fragen auftreten, die nur mit der Unterstützung von Kollegen gelöst werden können. Über die Expertensuche sollen Fragen wie „Gibt es jemanden in meiner Nähe, der dieses Problem bereits bearbeitet hat?“ und „Wie kann ich den zuständigen Experten am besten kontaktieren?“, „Wo befindet er sich gerade?“ beantwortet werden.

4.3 Benutzer & Zugriff

Die Zielgruppe von U-Know setzt sich aus den Mitarbeitern der Wirtschaftswissenschaftlichen Fakultät der Universität Regensburg, also Professoren, wissenschaftlichen Mitarbeitern und Verwaltungsangestellten, zusammen. An der Fakultät betreuen 21 Lehrstühle, die in vier Institute aufgeteilt sind, mit insgesamt ca. 80-100 wissenschaftlichen Mitarbeitern und Angestellten derzeit ca. 2.900 Studierende. Der Zugriff auf das System soll für WAP-fähige Mobiltelefone und PDAs mit WLAN-Anschluss realisiert werden.

Über WAP ist eine Ortsunabhängigkeit innerhalb der Abdeckung des GSM-Netzes sichergestellt. Ein sogenanntes WAP-Gateway (betrieben vom Netzanbieter) ermöglicht den Systemzugriff vom Mobiltelefon aus. Es wandelt Abfragen in HTTP-Abfragen um, welche an einen HTTP-Server weitergeleitet, dort bearbeitet, anschließend wieder an das WAP-Gateway und von dort aus an das Mobiltelefon zurückgegeben werden. Damit können Mitarbeiter (bei Netzverfügbarkeit) praktisch von überall auf Informationen wie z.B. auf das Mitarbeiterverzeichnis oder die Telefonnummern der Universitätsverwaltung zugreifen.

WLAN (Wireless Local Area Network) ermöglicht eine lokale Mobilität. Bei einem WLAN-Netz nach dem IEEE 802.11 Standard handelt es sich um ein

aus einzelnen Zellen bestehendes Funknetz, deren Reichweite je nach technischen Gegebenheiten auf 50-300m beschränkt ist. Geräte wie Notebooks oder auch PDAs, die mit einer Funkkarte ausgestattet sind, können über auf dem Gelände der Uni Regensburg verteilte Accesspoints eine drahtlose Verbindung ins Internet und damit zu U-Know aufbauen. Die schnelle Verbindung (bis zu 11Mbit/s), größere Displays und die bequemere Texteingabe über Tastatur oder Touchscreen ermöglichen eine komfortable Nutzung von Recherchefunktionen, den Zugriff und das Bearbeiten von Emails sowie die Darstellung längerer Dokumente. Für Besprechungen muss in der Regel mindestens einer der Beteiligten seinen Arbeitsplatz verlassen. Seinen Informationsbedarf kann er üblicherweise im Voraus nicht vollständig abschätzen und er hat nicht die Möglichkeit, sämtliche Informationsressourcen mit sich zu führen [Berg99]. Aus diesem Grund ist es sinnvoll, wenn dieser Teilnehmer mithilfe eines mobilen WMS am Ort der Zusammenkunft oder von unterwegs zumindest auf häufig benötigtes Wissen, wie etwa Ansprechpartner, oder interne und externe Informationsquellen zugreifen kann, ohne an stationäre Ressourcen angewiesen zu sein. Für die Benutzer bedeutet dies, dass die U-Know-Funktionen das Problem der begrenzten Informationsverfügbarkeit und -verarbeitungskapazität (s. Kapitel 3) zumindest abschwächen können.

5 Zusammenfassung und Bewertung

Die Herausforderungen der Zukunft, steigender Wettbewerb und Internationalisierung, werden in absehbarer Zukunft die Hochschulen ebenso betreffen wie die Unternehmen. In diesem Umfeld kommt dem vorhandenen Verwaltungswissen bzw. dem Wissen, wo überhaupt Informationen und Ansprechpartner zu finden sind, eine bedeutende Rolle zu. WMS unterstützen durch ihre Funktionalität den Zugriff auf die benötigten Informationen. Viele Mitarbeiter, die an der Universität tätig sind, verbringen jedoch einen beträchtlichen Teil ihrer Arbeitszeit nicht am Arbeitsplatz und können daher nicht auf die stationär gebundenen Informationsressourcen zugreifen. Deshalb wird an der Universität Regensburg der Prototyp U-Know als mobiles WMS entwickelt. Durch den Einsatz mobiler IuK-Technologien soll seine Nutzung erleichtert werden, da durch die zentrale Erfassung der häufig benötigten Informationen ein zeitaufwendiges Durchsuchen verschiedenen Server mit jeweils unterschiedlichen Navigationsstrukturen wegfällt und durch die Ausrichtung auf Mobiltelefone und PDAs eine größere Ortsunabhängigkeit erreicht wird.

Mitarbeiter der Wirtschaftswissenschaftlichen Fakultät können so flexibel auf häufig benötigte Informationen zugreifen. Arbeitsabläufe und damit For-

schung und Lehre gewinnen dadurch an Effizienz. Als Problem für eine Beurteilung nach den betriebswirtschaftlichen Kriterien Kosten und Ertrag wird sich dabei wohl die Tatsache erweisen, dass zwar die Kosten der mobilen Lösungen in Form von Hardware-, Software- und Betriebskosten genau ermittel- und zurechenbar sind, sich die Effizienzvorteile allerdings nur schwer quantifizieren lassen. Vor diesem Hintergrund gilt es, die Nutzung des Prototypen nach seiner Inbetriebnahme einer sorgfältigen Untersuchung zu unterziehen. Dadurch können die „Alltagstauglichkeit“ und die Vor- und Nachteile einer derartigen Lösung ermittelt werden und darüber hinaus weitere Benutzerwünsche, die sich im täglichen Gebrauch ergeben, in die zukünftige Entwicklung einfließen.

Prinzipiell kann man wohl davon ausgehen, dass die Bedeutung des Wissensmanagements generell und die Anforderungen an Kommunikation und Vernetzung von verteilten Mitarbeitern auch in Zukunft weiter zunehmen. Die Nutzung mobiler IuK-Technologien bietet hier ein besonderes Potenzial, um einen echten Mehrwert für „traditionelle“ Wissensmanagement-Lösungen zu schaffen. Mit den hier vorgestellten Überlegungen für einen mobil verfügbaren Prototypen wird ein erster Versuch in diese Richtung unternommen, der in weiteren Arbeiten zu einem umfassenden Konzept für mobiles Wissensmanagement ausgebaut werden soll.

6 Literatur

- [BaÖs99] Bach V., Österle H.: Wissensmanagement eine unternehmerische Perspektive. In: Bach, V.; Vogler P.; Österle, H. (Hrsg.): Business Knowledge Management, Praxiserfahrungen mit Intranet-basierten Lösungen. Springer Verlag, Berlin 1999.
- [Baye98] Bayerisches Staatsministerium für Wissenschaft, Forschung und Kunst (Hrsg.): Bayerisches Hochschulgesetz. München 1998.
- [BeBl96] Belotti, V.; Bly, S.: Walking Away from the Desktop Computer: Distributed Collaboration and Mobility in a Product Design Team. In: Ehrlich, K.; Schmandt, C. (Hrsg.): Proceedings of CSCW'96. ACM Press, Boston, Mass. 1996, S. 209-218.
- [Berg99] Bergqvist, Jens et. al.: Moving Out of the Meeting Room – Exploring support for mobile meetings. In: Bødker Susanne et. al. (Hrsg.): Proceedings of the Sixth European Conference on Computer Supported Cooperative Work (ECSCW'99). Kopenhagen, Dänemark 1999, S. 81-98.
- [Bind01] Binder, L.: „Mobile e-business“ – eine Erfolgsgeschichte? Download: [http://www.competencesite.de/mbusiness.nsf/A6BBC535836FE23DC1256A7900377F96/\\$File/mobilesebusiness_erfolgsgeschichte.pdf](http://www.competencesite.de/mbusiness.nsf/A6BBC535836FE23DC1256A7900377F96/$File/mobilesebusiness_erfolgsgeschichte.pdf), Abruf am 2002-03-07.
- [DiBe96] Dix, A.; Beale, R. (Hrsg.): Remote Cooperation: CSCW Issues for mobile and teleworkers. Springer, London 1996.

- [Fagr99] Fagrell, Henrik et al.: Exploring Support for Knowledge Management in Mobile Work. In: Bødker, S.; Kyng, M.; Schmidt, K. (Hrsg.): Proceedings of ECSCW'99, Klüwer, ORT 1999, S. 277-290.
- [Froh94] Frohlich D. et. al.: Informal Workplace Communication: What is it Like and How Might We Support it? Working Paper Personal Systems Laboratory HP Laboratories, Hewelett-Packard, Bristol 1994.
- [Hela99] Helal, Abdelsalam et al.: Any Time, Anywhere Computing, Mobile Computing Concepts and Technology. Kluwer Academic Publishers, Boston 1999.
- [Klos01] Klosa, O.: Wissensmanagementsysteme in Unternehmen. State-of-the-Art ihres Einsatzes. DUV, Wiesbaden 2001.
- [KrLj98] Kristoffersen, S.; Ljungberg, F.: Mobile Informatics. Innovation of IT use in mobile settings. IRIS '21 Workshop Report, Denmark 1998.
- [Lehn00] Lehner, F.: Organisational Memory. Konzepte und Systeme für das organisatorische Lernen und das Wissensmanagement. Carl Hanser Verlag, München 2000.
- [LeH+95] Lehner, F. et al.: Wirtschaftsinformatik. Theoretische Grundlagen. Carl Hanser Verlag, München 1995.
- [Maie02] Maier, R.: Knowledge Management Systems. Springer Verlag, Berlin 2002.
- [McDac98] McDonald, David, W.; Ackerman Mark, S.: Just Talk to Me: A Field Study of Expertise Location. In: Poltrock, S.; Grudin, J. (Hrsg.): Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work (CSCW'98), ACM Press, Seattle 1998, S. 315-324.
- [Mono02] Monopolkommission (Hrsg.): Wettbewerb als Leitbild für die Hochschulpolitik. Sondergutachten der Monopolkommission gemäß § 44 Abs. 1 Satz 4 GWB. Bonn 2000.
- [Perr01] Perry, M. et al.: Dealing with Mobility: Understanding access anytime, anywhere. In: ACM Transactions on computer human interaction (TOCHI) 8 (2001) 4, S. 323-347.
- [Schü96] Schüppel, J.: Wissensmanagement. Organisatorisches Lernen im Spannungsfeld von Wissens- und Lernbarrierern. DUV, Wiesbaden 1996.
- [Schu99] Schulte, B. A.: Organisation mobiler Arbeit, Der Einfluss von IuK-Technologien. DUV, Wiesbaden 1999.
- [Stei02] Stein, A.: Wissensmanagement - nicht nur eine organisatorische Aufgabe. Download unter: <http://www.bva.bund.de/aufgaben/win/beitraege/00046/>, Abruf am: 2002-05-13.



Die Wortschatz-Methode als netzbasierte Kreativitätstechnik in einer virtuellen Lernumgebung

Möglichkeiten und erste Erfahrungen

Gerhard Heyer¹, Thomas Hoppe¹, Katrin Müller²

¹Universität Leipzig
Institut für Informatik
Abt. Automatische Sprachverarbeitung
PF 920
D-04009 Leipzig
Germany

{heyer, hoppe}@informatik.uni-leipzig.de

²Fraunhofer-Institut
Arbeitswirtschaft und Organisation
Nobelstrasse 12
D-70569 Stuttgart
Germany

Katrin.Mueller@iao.fhg.de

Zusammenfassung

Die vorgestellte Kreativitätstechnik Wortschatz-Methode unterstützt mit automatisch generierten Assoziationen und Begriffen die Erarbeitung eines Themas. Sie lässt sich in unterschiedlichen betrieblichen und universitären Zusammenhängen, z.B. beim netzbasierten Lernen, sinnvoll zur Kreativitätsförderung einsetzen. Als Informationsquelle dient die Datenbank „Deutscher Wortschatz“. Es wird dargestellt, wie die Methode funktioniert und in welcher Weise sie netzbasiert genutzt werden kann.

Abstract

The presented creative method “Wortschatz-Methode” supports with automatically generated associations and terms the development of a topic. It works in a virtual learning environment and can be used in different connections meaningfully for creativity promotion. The data base “Deutscher Wortschatz” serves as source of information. We describe the surface and the functions of this new method.



1 Einleitung

In der Hochschullehre werden die Weichen neu gestellt. Zum einen wird verstärkt Wert gelegt auf die sinnvolle Vermittlung von Lehrinhalten mit Hilfe der neuen Medien, zum anderen wird darauf geachtet, dass ergänzend im Bereich „Schlüsselqualifikationen“ Inhalte vermittelt werden. Insbesondere Inhalte zur „Arbeitsmethodik“ sollen zukünftig nicht mehr am Rande der Ausbildung behandelt werden bzw. der individuellen Initiative überlassen werden.

Mit diesen Weichenstellungen wird selbstbestimmtes Lernen medial und arbeitsmethodisch unterstützt. Dem einzelnen Lernenden wird erlaubt - entsprechend seiner persönlichen Anlagen und Neigungen sowie seines aktuellen Lernbedarfs - Inhalte, Lernorte und Lernzeiten interaktiv selbst zu bestimmen oder auch im Team gemeinsam zu arbeiten [ELR00].

2 Grundlagen

Das Projekt „Moderation VR“ („Moderations- und Kreativitätsmodule in VR-Umgebungen“, gefördert vom BMBF, Projektträger Neue Medien in der Bildung und Fachinformation) zielt darauf ab, Moderations- und Kreativitätstechniken, die als psychologisch fundierte Arbeitsmethoden in der Praxis erprobt sind, elektronisch zu realisieren und für einen netzbasierten Einsatz in der Lehre nutzbar zu machen. Neben allgemein bekannten Verfahren wie Brainstorming, Mind Mapping, Kartenabfrage oder Reizwortanalyse wird im Folgenden speziell die im Rahmen des Projekts Moderation VR entwickelte „Wortschatz-Methode“ beschrieben. Diese neue Kreativitätstechnik beruht auf der Nutzung von Wortähnlichkeiten bzw. Wortassoziationen, die auf der Grundlage einer großen Wortdatenbank für die deutsche Sprache berechnet worden sind.

Hintergrund der Wortdatenbank ist das Projekt Deutscher Wortschatz am Lehrstuhl für Automatische Sprachverarbeitung der Universität Leipzig. Durch die Analyse sehr großer Textdatenbestände von offline und online verfügbaren Quellen ist in den vergangenen Jahren eine der größten Online-Datenbanken zum Deutschen Sprachgebrauch aufgebaut worden, mit vielfältigen Anwendungen u.a. in den Gebieten Web Searching, Software-Modellierung, Terminologieextraktion und adaptierbare News-Agentensysteme [He00],[He01].

Aufbauend auf der Erfassung und Analyse der Ausgangstexte werden statistische Auswertungsmethoden vor allem für die Ermittlung signifikanter Kollokationen eingesetzt. Die Methode kann somit zur automatisch generierten Ideensammlung genutzt werden. Dabei wird als Kontexteinheit entweder ein Satz („Satzkollokationen“) oder die unmittelbare Nachbarschaft zweier Muster (Wortformen) gewählt („Nachbarschaftskollokationen“). Neben der Darstellung von Kollokationsmengen als Begriffslisten erfolgt auch eine (real time-)Visualisierung als Begriffsnetzwerk, in dem die „stärksten“, untereinander verbundenen Kollokationen eines Ausgangsbegriffs dargestellt werden (vgl. Abb. 2).

3 Wortschatz-Methode

Die Datenbank „Deutscher Wortschatz“ enthält neben den aufgenommenen Wörtern auch ihre Beziehungen untereinander. Der Hauptvorteil der Wortschatz-Abfrage liegt in der schnellen Ermittlung der in Abb. 2 bis 5 dargestellten Visualisierungen und Kollokationen. Die bisherigen Auswertungen des Datenbankbetriebs ergaben eine zufriedenstellende Performance. In der Wortdatenbank liegen umfangreiche Datenbestände auch aus anderen Forschungsbereichen vor, die sich über geeignete Abfragen zu vielfältigen Informationsausgaben variabel zusammenstellen lassen und dadurch ergänzende Aussagen zu einem gewählten Suchbegriff ermöglichen.



Abb. 1: Situation „Wortschatz-Methode“

Am Beispiel-Thema „Projektplanung für eine Softwareentwicklung“ soll im Folgenden das konkrete Vorgehen erläutert werden.

In der ersten Phase wird eine Skizze z.B. eines ersten Pflichtenhefts zusammengestellt. Nach der etwaigen Beschaffung von Detailinformationen kann der Entwurf weiter präzisiert werden.

Danach kommt die zweite Phase, bei der die Wortschatz-Methode angewendet wird. Nach der Eingabe des Suchbegriffs „Projekt“ werden andere Worte (Kollokationen), die in signifikanten Beziehungen zum Suchbegriff stehen, automatisch generiert.

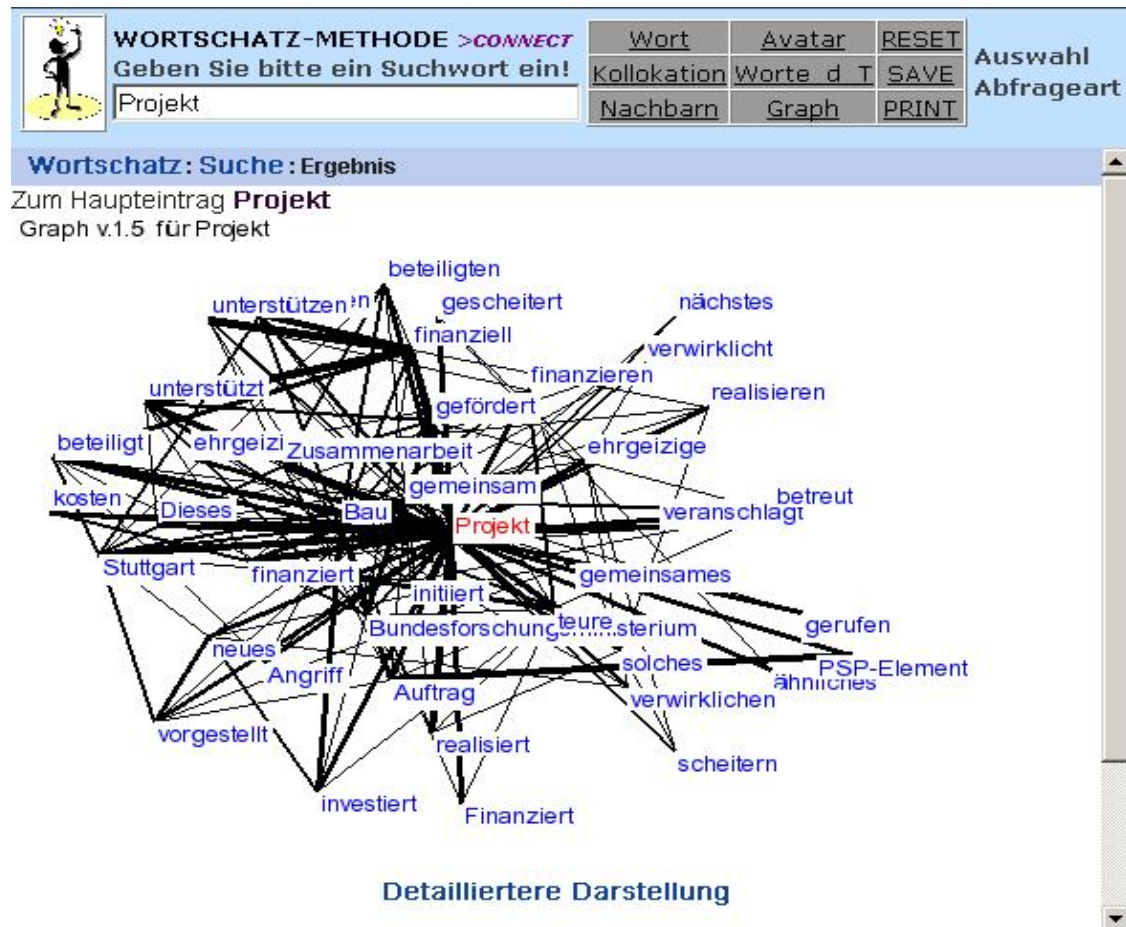


Abb. 2: Graph zum Suchbegriff „Projekt“

Diese Wörter bewirken nun, dass Assoziationen (neue Gesichtspunkte, Begriffe, Zusammenhänge usw.) aufgedeckt werden oder auch das bereits Bekannte bestätigt wird.



Abb. 3: Ausschnitt Angaben zum Wort „Projekt“

zum beschriebenen Vorgehen können die gefundenen Begriffe für eine spätere Verwendung gespeichert werden.

Neben der Betriebsart „Abfrage mit Suchbegriff“ kann ein Dialog mit einem Avatar geführt werden, bei dem abwechselnd eine Frage eingegeben wird und der Avatar antwortet.

WORTSCHATZ-METHODE >CONNECT
Geben Sie bitte ein Suchwort ein!

Wort	Avatar	RESET
Kollokation	Worte d T	SAVE
Nachbarn	Graph	PRINT

Auswahl
Abfrageart

Fragen Sie bitte unseren Avatar:



Mein Name ist Wortschatz-Expertin (dbBot)

> CONNECT
"Hallo hier vom Avatar und vielen Dank fuer den Aufruf!"

A.L.I.C.E.
ARTIFICIAL LANGUAGE INTERPRETER COMPUTER SYSTEM
Powered by AIML (Artificial Intelligence Markup Language)

Abb. 6: Wortschatz-Methode mit Avatar


Er nutzt dazu Informationen der Wortschatz-Datenbank wie Kollokationen, Beispielsätze, Sachgebiete, Beschreibungen und weitere statistische und grammatikalische Angaben.

WORTSCHATZ-METHODE >CONNECT
Geben Sie bitte ein Suchwort ein!

Wort	Avatar	RESET
Kollokation	Worte d T	SAVE
Nachbarn	Graph	PRINT

Auswahl
Abfrageart

Fragen Sie bitte unseren Avatar:



Mein Name ist Wortschatz-Expertin (dbBot)

> Welches Sachgebiet ist Projekt
"Das Wort laesst sich zu folgenden Sachgebieten zuordnen: Ökonomie, Nachname, Allgemeines/Interdisziplinäre/Allgemeinwörter, Rechnungswesen/Steuer/Revision, Allgemeines/Unternehmen/Management und Tätigkeiten ."

A.L.I.C.E.
ARTIFICIAL LANGUAGE INTERPRETER COMPUTER SYSTEM
Powered by AIML (Artificial Intelligence Markup Language)

Abb. 7: Wortschatz-Methode - Abfrage mit Avatar

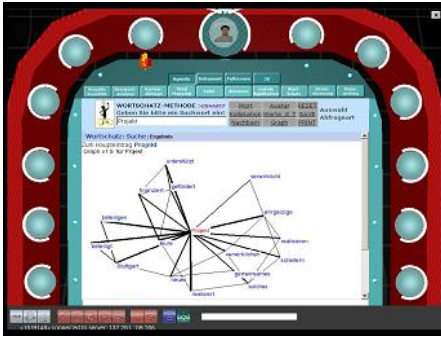


Abb. 8: Virtuelle Arbeitsumgebung mit „Wortschatz-Methode“



Abb. 9: Virtuelle Arbeitsumgebung in 3D-Darstellung

Als technische Plattform wird die e-Learning-Suite von Hyperwave genutzt, die um Macromedia-Director-Komponenten ergänzt wurde und somit eine virtuelle 3D-Multiuser-Umgebung für das Modul Wortschatz-Methode liefert (Abb. 8, 9). Die Arbeit mit dem Wortschatz-Modul und die Zugriffe auf die Wortschatz-Datenbank werden über Internet-Verbindungen realisiert.



Abb. 10: E-Learning-Suite von Hyperwave - Grobstruktur

Zur Einführung in die Methode steht innerhalb der Lernumgebung ein Modul mit Beschreibung, Hinweisen und Beispielen bereit. Bei Tests mit unterschiedlichen Nutzergruppen wurde die Kreativitätsmethode mit Interesse aufgenommen und als sehr anregend empfunden. Gegenwärtig laufende umfassende Tests bestätigen diese Ergebnisse.

4 Resümee

Die beschriebene Wortschatz-Methode fördert die Kreativität beim Herangehen an Informatik-Themen, bietet Lernhilfen und unterstützt beim Ideenfinden, indem sie auf benachbarte Zusammenhänge hinweist. Es können neue

Begriffswelten in der Informatik, sowie anderen Fachgebieten, aber auch beim Sprachenlernen erschlossen werden. Auch ist die Methode zur Rechercheunterstützung nutzbar, indem aus abgefragten relevanten Kollokationen Suchbegriffe ausgewählt und zusammengestellt werden.

5 Literatur

- [He00] Heyer, G.; Läuter, M.; Quasthoff, U.; Wolff, Chr. "Texttechnologische Anwendungen für Inter- und Intranet". In: Schmitz, Klaus-Dirk (ed.) (2000). Sprachtechnologie für eine dynamische Wirtschaft im Medienzeitalter. Tagungsakten der XXVI. Jahrestagung der Internationalen Vereinigung Sprache und Wirtschaft, Köln, November 2000. Wien: TermNet Publisher, 203-209.
- [He01] Heyer, G.; Läuter, M.; Quasthoff, U.; Wittig, Th.; Wolff, Chr. "Learning Relations using Collocations" In: A. Maedche, S. Staab, C. Nedellec and E. Hovy, (eds.). Proc. IJCAI Workshop on Ontology Learning, Seattle/ WA, August 2001, 19 - 24.ö
- [ELR00] Encarnacao, J.; Leidhold, W.; Reuter, A. "Szenario: Die Universität im Jahre 2005" Z. Informatik Spektrum 23 (2000) 4, S. 264 - 270. Springer-Verlag Berlin, Heidelberg 2000



Modellbildung und Visualisierung: Das Sowinet.de-Wahlmodell

Maximilian Eibl¹, Markus Quandt²

¹GESIS – IZ Sozialwissenschaften
Schiffbauerdamm 19
10117 Berlin

eibl@berlin.iz-soz.de

²GESIS – Zentralarchiv für
empirische Sozialforschung
Bachemer Str. 40
50931 Köln

markus.quandt@uni-koeln.de

Zusammenfassung

Informationsvisualisierung wird in der Regel in drei unterschiedlichen Einsatzarten genutzt: Datenanalyse, Navigation und Wissensvermittlung. Dabei entstehen in Bezug auf den jeweiligen Einsatzbereich und unabhängig vom inhaltlichen Anwendungsfeld ganz unterschiedliche Formen der Visualisierung. Dieser Artikel stellt eine Möglichkeit vor, diese drei Verwendungskontexte von Informationsvisualisierung in einer Darstellung zu integrieren. Den Anwendungskontext dazu bildet ein sozialwissenschaftliches Modell des individuellen Wahlverhaltens bei politischen Wahlen, dessen Umsetzung im Rahmen des Projekts „sowinet“ hier vorgestellt wird.

Abstract

Information visualization is usually used in three areas of application: data analysis, navigation, and knowledge transfer. Here, totally different kinds of visualizations are developed. This article describes an example that integrates the three areas into one single graphical representation. The application context is a sociological model of voting.

1 Einleitung

In diesem Artikel wird die Visualisierung eines Modells vorgestellt, das die Einflüsse auf das individuelle Verhalten bei politischen Wahlen widerspiegelt. Diese Visualisierung wurde im Rahmen des Projekts „sowinet.de“ erstellt. Bei „sowinet.de“ handelt es sich um ein online-Magazin, das sozialwis-



senschaftliche Themen und Fragestellungen populärwissenschaftlich aufbereitet im Internet präsentiert. Es soll dabei ein Spektrum von Zielgruppen angesprochen werden, das von Schülern bzw. Lehrern über Journalisten zu allgemein interessierten Bürgern reicht. Das Magazin ist unter <http://www.sowinet.de/> einzusehen.

Das Projekt lief vom April 2000 bis März 2002 und wurde vom Bundesministerium für Bildung und Forschung (BMBF) gefördert. Die inhaltliche Thementauswahl und -bearbeitung erfolgte durch das Zentralarchiv für empirische Sozialforschung, Köln, in Zusammenarbeit mit Journalisten. Die Hochschule für Gestaltung, Offenbach, war verantwortlich für das ästhetische Erscheinungsbild. Und das Informationszentrum Sozialwissenschaften in Bonn und Berlin führte die technische Umsetzung durch. Einen Überblick über die Gesamtkonzeption und Umsetzung dieses interdisziplinären Vorhabens geben Bürdek et al. 2001, Eibl 2001 und Quandt 2001.

Der vorliegende Beitrag fokussiert auf ein Einzelelement von „sowinet.de“: das Wahlmodell. Der Anwendungskontext ist das individuelle Wahlverhalten bei politischen Wahlen. Hierbei handelt es sich um eine in der Wahlforschung weitgehend akzeptierte und belegte Theorie. Aufgrund der alltagsweltlichen Anschlussmöglichkeiten eignet sich das Modell gut für den Einbau in ein populärwissenschaftliches Format. Aufgrund der Komplexität des Modells, das aus zahlreichen untereinander wirkenden Elementen verschiedener Ordnung besteht, bietet sich wiederum ein gewinnbringender Einsatz von Visualisierung an. Im Folgenden werden nun der theoretische Hintergrund, die Konzeption und Umsetzung dieses Modells beschrieben.

2 Aspekte der Visualisierung

Methoden der Informationsvisualisierung werden in den verschiedensten Verwendungskontexten eingesetzt. Drei grobe Bereiche lassen sich dabei herauskristallisieren: Datenanalyse, Navigation und Wissensvermittlung. Dabei entstehen in bezug auf den jeweiligen Verwendungszweck und unabhängig vom inhaltlichen Anwendungsfeld ganz unterschiedliche Formen der Visualisierung.

Der wohl am ausgiebigsten erforschte Kontext ist die Datenanalyse. Hier erlauben bereits einfache graphische Darstellungen in Form von Diagrammen Aussagen über generelle Trends, Muster oder Besonderheiten (für den sozialwissenschaftlichen Kontext siehe grundlegend Schnell 1994, Tufte 1983). Komplexere Anwendungen lassen eine direkte Manipulation solcher Dia-

gramme zu. So beschreibt beispielsweise Wolff 1996 ein Werkzeug zur Faktenrecherche, das graphische Eingabemöglichkeiten in Diagrammen erlaubt. Weniger traditionelle Darstellungen zur Analyse stellen Visualisierungen im Document Retrieval zur Verfügung. Die Einsatzmöglichkeiten sind hier sowohl was die Form der Darstellung, als auch was den Zeitpunkt innerhalb des Rechercheprozesses angeht, zahlreich. So nutzt beispielsweise LyberWorld eine dreidimensionale Baumstruktur für eine dem eigentlichen Retrievalprozess vorgeschaltete Termselektion (Hemmje et al. 1994). Verschiedene auf Venn-Diagrammen oder Hypergraphs basierende Visualisierungen versuchen, die Anfrage, bzw. die Anfragesprache SQL graphisch aufzubereiten (z.B.: Michard 1982). Systeme wie InfoCrystal (Spoerri 1994) oder DEViD (Eibl 2000 und Eibl 2002) versuchen, Anfrage und Ergebnis graphisch zu integrieren. Andere Systeme wiederum visualisieren die Ergebnismenge als zwei- oder dreidimensionale Dokument-Term Beziehungsmuster (z.B.: Korfhage 1991; Elzer & Krohn 1999). Ein sehr detaillierter Ansatz zeigt die Vorkommnisse der Suchkriterien in Dokumenten in Form spezieller TileBars (Hearst & Peterson 1996) oder verkleinerten Darstellungen der Dokumente (Odgen et al. 1998). Da die Nützlichkeit solcher Visualisierungen jedoch oft genug auf singuläre Fragestellungen zu einem bestimmten Zeitpunkt im Retrievalprozess beschränkt ist, versucht INSYDER (Reiterer 2001, Klein et al. 2002) ihre Wirkung durch die Integration verschiedener Visualisierungen innerhalb eines intuitiv erfassbaren Gesamtkonzepts zu erhöhen. Eine zusammenfassende Übersicht zu den verschiedenen Ansätzen und ihren konzeptionellen Hintergründen bieten Eibl 2000 und Eibl 2002.

Ein anderer Anwendungskontext ist die Navigation in Datenbeständen. Hier werden in der Regel Baumstrukturen zur Navigation in hierarchischen Datenbeständen herangezogen. Verschiedene Methoden der Verzerrung oder des Zusammenklappens von Teilästen erlauben interessante interaktive Anwendungen. Für einen Überblick siehe Roppel 1996. Im World Wide Web (WWW) werden solche Visualisierungen gerne als sogenannte Sitemaps eingesetzt, die dem Leser einen Überblick über Struktur und Inhalt der Site geben sollen. Däßler 1998 stellt ein VRML-basiertes System zur Visualisierung hierarchischer Strukturen vor. Darüber hinaus finden sich auch verschiedene auf der Topographie-Metapher basierende Visualisierungen (siehe Eibl & Mandl 2002a). Hier werden Dokumente im zwei- oder dreidimensionalen Raum so angeordnet, dass die Entfernung zwischen den einzelnen Dokumenten ihrer Ähnlichkeit entspricht. Semantische Nähe wird also durch räumliche Nähe ausgedrückt. Dabei entstehen an vereinzelter Stellen Häufungen, die in Form Topographischer Karten visualisiert werden können. Typisch für die Erstellung solcher Karten ist die Unvorhersehbarkeit der Darstellung, die dar-

aus herrührt, dass verschiedene Verfahren zu verschiedenen Ergebnissen führen (Eibl & Mandl 2002b).

Schließlich eignet sich Informationsvisualisierung gut zur Wissensvermittlung. Hier ist vor allem die graphische Darstellung von Modellen sehr hilfreich. Solche Modelle können verschiedenster Natur sein. Im einfachsten Fall sind es eins-zu-eins Kopien natürlicher Vorbilder, beispielsweise ein für den Biologieunterricht erstelltes animiertes VRML-Modell einer Hummel, welches das Flugverhalten abbildet. In einem komplexeren Fall können auch graphische Darstellungen von nicht sichtbaren Objekten und Phänomenen eingesetzt werden. Berichte über Erkenntnisse der Weltraumforschung kommen ohne solche Darstellungen nicht aus. Schließlich können graphische Modelle noch dynamische Prozesse wie beispielsweise das Räuber-Beute-Phänomen darstellen.

Je abstrakter der Gegenstand der Modellierung ist, desto stärker kommen die Vorteile computergestützter Visualisierungen zum Tragen. Sie erlauben nämlich Anschaulichkeit, ohne den modellhaften Charakter der Visualisierung zu verbergen (Reiners & Saborowski 1997). Damit sind sie besonders geeignet, wenn die Modellhaftigkeit des visualisierten Gegenstands selbst Teil der zu vermittelnden Botschaft ist. Sozialwissenschaftliche Theorien sind hierfür ein Musterbeispiel.

3 Modelle der Wahlforschung

Ein thematischer Schwerpunkt von „sowinet.de“ ist die Darstellung von Theorien, Daten und Forschungsergebnissen aus der Wahlforschung. Dabei ist im Kontrast zu der ansonsten kurzfristig orientierten Berichterstattung des Magazins, die thematisch nur durch aktuelle Ereignisse und Forschungsergebnisse gebunden ist, ein vertiefender Einblick in einen abgegrenzten Forschungsbereich zu geben. Das Themenfeld Wahlforschung ist jedoch vielschichtig, sowohl die Anzahl und Breite der möglichen Forschungsfragen betreffend als auch von der Vielfalt der wissenschaftlichen Perspektiven her.

Das heißt, dass nicht nur eine einzelne Theorie darzustellen ist, sondern dass den Nutzern eine gewisse Orientierung in dem ganzen Themenfeld gegeben werden soll. Zugleich soll aber auch ein Grundverständnis für die hier typischen wissenschaftlichen Denkweisen sowie für einzelne Argumentationsstränge vermittelt werden. Demnach sind nicht Stringenz und formale Eleganz die Hauptziele der visuellen Darstellung, wie sonst bei Wissenschaftsgrafiken üblich, sondern es sind zwei Ziele zu verfolgen, die leicht in Konflikt geraten

können: Erstens muss die Visualisierung flexibel genug sein, um inhaltlich heterogene Elemente und Sichtweisen zu integrieren. Zweitens muss aber in jeder Perspektive die Anschaulichkeit gewahrt bleiben, da von dem angesprochenen Nutzerkreis kaum die Bereitschaft zu größeren und wiederholten Abstraktionsleistungen erwartet werden kann.

Um diese Problematik zu verdeutlichen, wird in den beiden folgenden Abschnitten ein Aufriss des theoretischen Modells gegeben.

3.1 Das zu visualisierende Hypothesensystem

Deterministische Erklärungen sind in der Wahlforschung – wie auch sonst in den Sozialwissenschaften – kaum möglich. Aus naturwissenschaftlichen Anwendungen bekannte Funktions- oder Verlaufsdiagramme sind somit oft irreführend, weil sie eine eher „mechanische“ Idee vom Phänomen geben. Dagegen zeichnen sich menschliches Verhalten und gesellschaftliche Prozesse durch den Umgang mit Informationsunsicherheit und durch Pfadabhängigkeit aus. Soziologische Modelle müssen dementsprechend auch Rauschen und Komplexität abbilden. Auch sind der Formalisierung von soziologischen Theorien durch die schwer überwindbare Unschärfe mancher Begriffe Grenzen gesetzt. Dennoch liegt der wissenschaftlichen Arbeit selbstverständlich eine Reihe von ausgearbeiteten Denkmodellen zu Grunde, wenn diese auch relativ offen und flexibel sind.

In der Wahlforschung ist ein Modell vorherrschend, welches das individuelle Wahlverhalten zum Ausgangspunkt nimmt (für eine Lehrbuchdarstellung vgl. z.B. Bürklin & Klein 1998). Da dieser Ansatzpunkt auch alltagsweltlich anschaulich ist, wurde die Gliederung der Inhalte auf diesem sog. sozialpsychologischen Modell aufgebaut. Ausgehend vom Verhalten einzelner Menschen wird dann der Horizont auf die engere soziale Umgebung erweitert, diese wiederum wird in ihrem institutionellen und gesellschaftlichen Umfeld eingebettet.

- Der Kern des Modells besteht aus einfachen psychologischen Annahmen über die Wirkung von Einstellungen zu Parteien, Kandidaten und politischen Streitfragen auf das individuelle Wahlverhalten. Einen Schritt vorgelagert sind Hypothesen über die Entstehung dieser kurzfristig relevanten Einstellungen aus weiteren, tiefer verankerten Einstellungen, persönlichen Werten und grundlegenden politischen Überzeugungen.
- Nun wird die „Innensicht“ des Wählers verlassen und seine Einbettung in die Umwelt untersucht. Merkmale wie das Geschlecht, die Konfessi-

onszugehörigkeit, Beruf, Einkommen und Bildung sind einerseits individuelle Eigenschaften des Wählers, andererseits beschreiben sie seine Position im sozialen Gefüge. Diese soziodemografischen Merkmale wirken also auf zwei Wegen prägend auf die vorher genannten Einstellungen: Sie lassen erstens Schlüsse auf typische Lebenslagen einer Person zu, auf ihre Erfahrungen und Ressourcen.

- Zweitens geben die soziodemografischen Merkmale Auskunft über die Zugehörigkeit zu sozialen Großgruppen wie Gewerkschaften, Kirchen, Ethnien etc., die möglicherweise an langfristig angelegten politischen Konflikten beteiligt sind.
- Großgruppen und ihre Konflikte können wiederum auch auf einer Ebene untersucht werden, die klar außerhalb des Individuums liegt. Hier sind die politischen Interessenkonstellationen in einer Gesellschaft zu berücksichtigen, die wirtschaftliche Lage, die typischen Eigenheiten der sozialen Milieus, Netze und Gruppen, das Wahlrecht usw.

3.2 Inhaltliche Ebenen des Anwendungsgebietes

Auf der inhaltlichen Seite überbrückt das theoretische Modell also mehrere analytische Ebenen: Ein intra-individuelles, psychologisches Modell eines Akteurs ist in ein Netzwerk der Wirkungen sozialstruktureller Merkmale eingebettet. Dieser Einbettungsprozess wiederum muss vor dem Hintergrund eines gesamtgesellschaftlichen Bildes verstanden werden, für das modellhafte Abstraktionen kaum existieren. Vielmehr geht es hier oft um die Beschreibung einmaliger historischer Konstellationen. Die visuelle Darstellung soll nun den Blick auf den gesamten politischen Prozess wahren, aber zugleich die Ebenenwechsel erkennbar machen. Dies ist allein deshalb notwendig, weil die Beziehungen zwischen den Elementen auf jeder Ebene spezifische Bedeutungen haben: Im psychologischen Kern des Modells stehen die Elemente überwiegend in einseitigen Kausalverknüpfungen. Im Kranz der soziodemografischen Merkmale wird ein Netz aus Wirkbeziehungen beschrieben, das von wechselseitigen Kausalverknüpfungen geprägt ist, gleichzeitig wird die Ebenenverortung unscharf: Man kann die Elemente dieser Ebene sowohl als Beschreibung des individuellen Wählers sehen wie auch als Beschreibung der Gruppen, denen er oder sie angehört. Im Feld der gesellschaftlichen und institutionellen Gegebenheiten schließlich geht es für die Wahlforschung meist um Hintergrundfaktoren, die selten selbst Forschungsgegenstand sind, die also aus der Modell-Perspektive eher "zufällig" nebeneinander stehen als systematisch miteinander zusammenhängen.

Perspektivenwechsel anderer Art ergeben sich durch die Unterscheidung zwischen empirischen und theoretischen Sichtweisen: Neben Modellierung und Erklärung besteht ein großer Teil der sozialwissenschaftlichen Arbeit - gerade in der Wahlforschung - in der Erfassung gesellschaftlicher Zustände und Veränderungen. Einige Ergebnisse dieser Tätigkeit sind auch Zeitungslesern wohlbekannt, werden Meinungsumfragen über Wahlabsichten etc. doch regelmäßig veröffentlicht.

Mit Hilfe des skizzierten theoretischen Rahmens ergibt sich für solche punktuell veröffentlichten Daten allerdings ein wesentlich tieferes Verständnis. Dazu ist eine Überblendung von abstrakter Theorie und den empirischen Konkretisierungen notwendig, um empirische Ergebnisse auf theoretische Konstrukte beziehen zu können und umgekehrt.

4 Konzeption der Visualisierung des Wahlmodells

Diese mehrdimensionale inhaltliche Struktur war in eine Form zu bringen, die gleichzeitig Lösungen für drei instrumentelle Teilaufgaben bietet: 1. Die Mitteilung zentraler Ergebnisse der Wahlforschung, d.h. den Zugriff auf Einzelinformationen, 2. die Ausleuchtung der thematischen Bandbreite und 3. die Darstellung der hinter diesen einzelnen Facetten liegenden theoretischen Struktur, also die Förderung von Tiefenverständnis und Lernen. Die äußere Form der Präsentation ließ sich zwanglos der Gesamterscheinung der „sowinet“-Site anpassen, indem die inhaltlichen Aspekte in kurzen Texten vorgestellt werden. Diese Texte erscheinen als je eigene Webseiten mit einer Länge von zwei bis drei Bildschirmseiten. Für die "innere" Organisation musste jedoch eine Struktur entworfen werden, die Laien die Inhalte technisch wie kognitiv zugänglich macht. Die Möglichkeiten des Mediums Internet erlauben es, diese innere Organisation auch visuell darzustellen und sogar als eine Art Navigationswerkzeug mit den Inhalten zu integrieren.

Die Visualisierung muss auf der Oberfläche zusammenfassend folgende Elemente berücksichtigen:

- **Internale Faktoren:** Faktoren, die direkt den Entscheidungsprozess des einzelnen Wählers bestimmen, wie Wertvorstellungen, Parteipräferenz, Weltanschauung, etc.
- **Umweltbedingungen:** Faktoren, die auf den einzelnen Wähler „von außen“ einwirken, wie die Wirtschaftslage, das Wahlsystem, Medien etc.

- Mischfaktoren: Eine Zwischenposition zwischen internalen Faktoren und der Umwelt nehmen die soziodemografischen Merkmale wie Konfession, Beruf, Bildung etc. ein.
- Wirkungsbeziehungen: Die einzelnen Faktoren können direkt auf das Wahlverhalten wirken, oder aber auch indirekt über andere Faktoren, wodurch eine Wirkungskette entsteht. So wirkt beispielsweise die Herkunft auf die politische Weltanschauung, diese auf die Kandidatenpräferenz und diese wiederum auf das individuelle Wahlverhalten.

4.1 Das Grundmodell

In einem ersten Schritt wurden verschiedene Möglichkeiten der visuellen Umsetzung dieser Elemente untersucht. Hierbei standen unterschiedlichste Varianten zur Diskussion. Letztendlich fiel die Entscheidung auf eine relativ schlichte Darstellung, die einfache Interpretations- und Manipulationsmöglichkeiten zulässt und gleichzeitig die inhaltliche Offenheit und Flexibilität des Modells zum Ausdruck bringt.

Das Erscheinungsbild des Grundmodells ist in Abbildung 1 dargestellt. Internale und Umwelt-Faktoren sind in zwei verschachtelten wolkenartigen Gebilden untergebracht. Ausschlaggebend für die Visualisierung als Wolken war die Prämisse, dass die einzelnen Faktoren nicht ortsfest wirken sollten. Ihre Lokalisierung soll die Zugehörigkeit zu den unterschiedlichen Ebenen verdeutlichen, aber die genaue Positionierung der einzelnen Elemente zueinander soll nicht als bedeutungstragend interpretiert werden.

Die Mischfaktoren liegen auf einer Grenzlinie zwischen den verschachtelten Wolken, wobei diese Linie gewissermaßen durch die Mischfaktoren gebildet wird. Dadurch wird verdeutlicht, dass diese Faktoren in beide Wolken hineinwirken bzw. Übergänge dazwischen darstellen.

Links neben dem eigentlichen Modell steht ein Textbereich zur Verfügung, der Beschreibungen zum Wahlmodell insgesamt beziehungsweise zu den einzelnen Faktoren enthält.

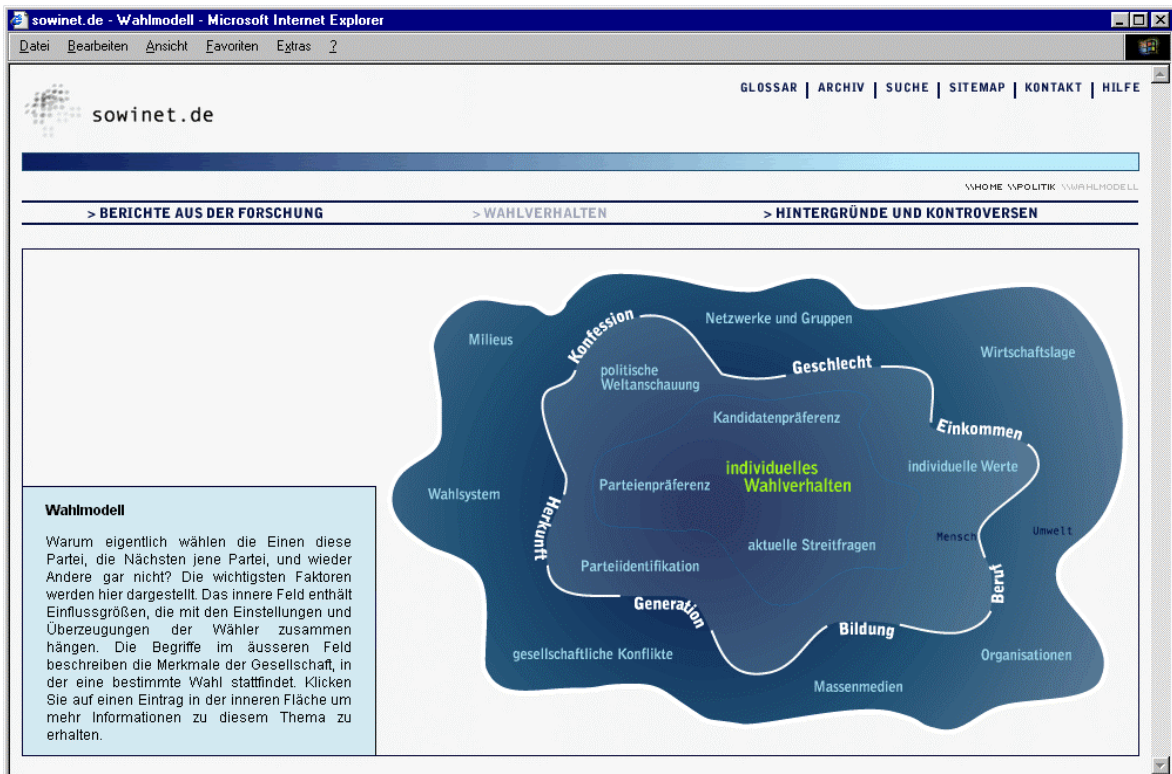


Abb. 1: Ausgangsansicht des Wahlmodells

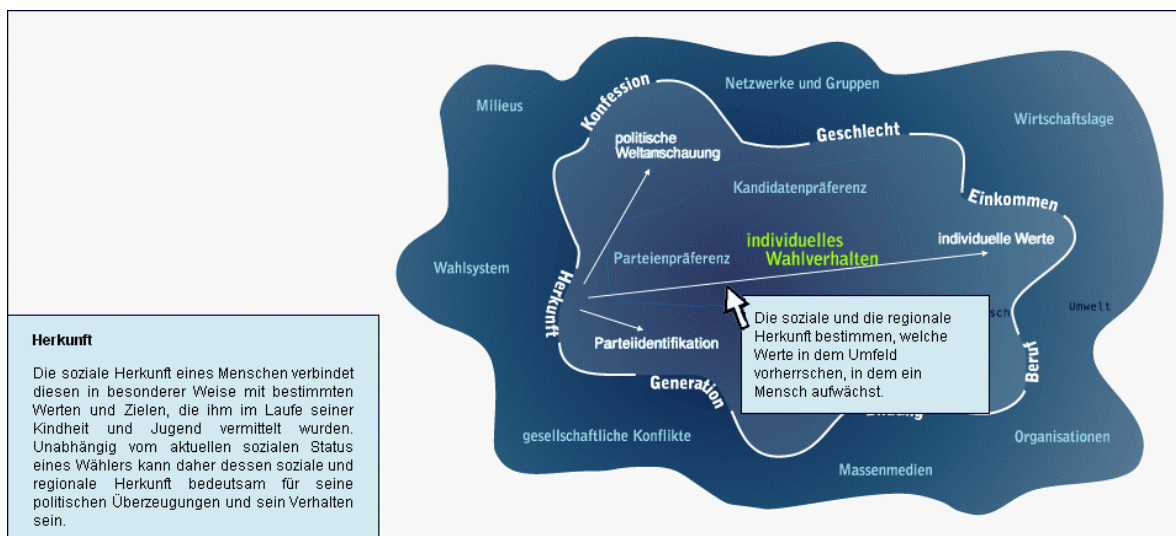


Abb. 2: Wahlmodell nach Klick auf das Element „Herkunft“: Drei Pfeile erscheinen, welche die Wirkungen der Herkunft auf andere Elemente verdeutlichen. Beim Überfahren der Pfeile mit der Maus erscheint eine kleine Informationsbox.

Die Faktoren des Wahlmodells sind anklickbar. Abbildung 2 zeigt die Aktivierung des Faktors „Herkunft“. Im linken Beschreibungsbereich erscheinen Informationen zum Faktor „Herkunft“. Zusätzlich erscheinen drei Pfeile, welche die Wirkung des Faktors auf „politische Weltanschauung“, „individuelle Werte“ und „Parteiidentifikation“ kennzeichnen. Beim Überfahren der Pfeile

mit der Maus erscheint eine kleine Informationsbox, die die Art der Wirkungen erläutert.

Diese Darstellung ermöglicht das schnelle, überblicksartige Verstehen des vergleichsweise komplexen Wahlmodells. Zugleich können auf der Basis der Darstellung mit Hilfe einfacher interaktiver Möglichkeiten zu einzelnen Elementen genauere Erklärungen gefunden werden. So kann bereits diese einfache Grundform durch die Einladung zu explorativem Erforschen ein relativ gutes Verständnis der Materie vermitteln.

4.2 Erweiterung 1: Inhaltliche Strukturierung und Linksammlung

In einem zweiten Schritt wurde das Wahlmodell um die Möglichkeit der Verlinkung zu anderen Informationsquellen angereichert. Zunächst war angedacht, das Wahlmodell als Navigationswerkzeug für die wahlpolitischen Seiten von „sowinet“ einzusetzen. Als solches eignet es sich allerdings nur eingeschränkt, da Wahlmodell und „sowinet“-Seiten aus Platzgründen nicht gleichzeitig auf dem Bildschirm präsentiert werden können. Durch die selektive Präsentation unterschiedlicher Verweise kann das Modell jedoch gut die kognitive Struktur der in „sowinet“ vorhandenen Texte ausdrücken.

Darüber hinaus werden zusätzliche Verlinkungen im Sinne einer Linksammlung angeboten. Dazu werden zu den jeweils aktivierten Elementen zusätzlich Verweise angezeigt. Die Gruppierung erfolgt nach zwei Kriterien. Zunächst werden zu den einzelnen Elementen jeweils nur die inhaltlich relevanten Links angezeigt. Dann sind die Links zusätzlich nach Herkunft geordnet: Links zu Artikeln von „sowinet“, Links zu Seiten der GESIS als Projektausführer und inhaltlich Verantwortlicher und schließlich Links zu anderen Internetressourcen.

Abbildung 3 zeigt die Links zur allgemeinen Beschreibung des Wahlmodells. Hier wurden gerade der Verweis „»sowinet-Links“ angeklickt. Es erscheinen fünf Verweise auf Artikel von „sowinet“.

Während das Grundmodell ein überblicksartiges Verständnis des Wahlverhaltens vermittelt, können durch die Erweiterung zu einer Linksammlung tiefergehende Informationen zu Einzelaspekten zur Verfügung gestellt werden. Durch die dezente Einbindung wird die Interpretation des Modells selbst aber keineswegs komplexer.



Abb. 3: Wahlmodell mit zusätzlichen Links: Hier wurde die Linkliste zu „sowinet-Links“ geöffnet.

4.3 Erweiterung 2: Wählergruppenanalyse

Die nächste Erweiterung demonstriert die Darstellung empirischer Ergebnisse innerhalb des Modellrahmens. Zunächst wurden mittels eines statistischen Verfahrens (näheres siehe 5) auf der Basis von Umfragedaten zur letzten Bundestagswahl neun herausgehobene Wählergruppen ermittelt. Diese sind:

1. SPD-Wähler aus der Arbeiterschicht, protestantisch oder konfessionslos
2. SPD-Wähler aus der neuen Mittelschicht, unter 45 Jahren
3. CDU-Wähler aus der neuen Mittelschicht, über 50, katholisch/kirchengebunden
4. CDU-Wähler unter kirchengebundenen Arbeitern
5. CDU-Wähler unter Landwirten und Selbständigen
6. FDP-Wähler
7. GRÜNEN-Wähler unter Angestellten und Beamten
8. GRÜNEN-Wähler unter Nicht-Erwerbstätigen
9. Nicht-Wähler

Mit Hilfe einer DropDown-Box im linken Bereich kann der Anwender nun spezifische Wählergruppen anwählen. Zu der selektierten Gruppe werden zusätzliche Informationen angezeigt:

- Der Prozentsatz der Stimmen, den die in dieser Gruppe dominante Partei bei der letzten Bundestagswahl *im Gesamtdurchschnitt* erhalten hat.

- Der Stimmenanteil, den die dominante Partei bei der letzten Bundestagswahl *in dieser Wählergruppe* erhalten hat. Dabei wird ein Unscharfbereich kenntlich gemacht, der der Anteilsschwankung in Untergruppen entspricht.
- Die Ausprägung der Faktoren, die für diese Wählergruppe typisch sind.

Abbildung 4 zeigt den Einsatz der Wählergruppenanalyse: Neben dem Wahlmodell erscheinen Informationen zur Gruppe der „CDU-Wähler unter Landwirten und Selbständigen“. Ein Balken zeigt graphisch, welchen Stimmenanteil die CDU insgesamt bei der Wahl erhalten hat (Dreiecksmarkierung) und wie viel Prozent der Stimmen sie im Kontrast in dieser speziellen Wählergruppe bekam. Im Begleittext wird zudem auf weitere Untergruppen verwiesen, außerdem sind wiederum verschiedene Links zur theoretischen Vertiefung zugänglich. Im Wahlmodell rechts werden Kernmerkmale (hier Konfession = katholisch, Beruf = Landwirt, Generation = vor 1960 geboren) der Gruppe hervorgehoben.

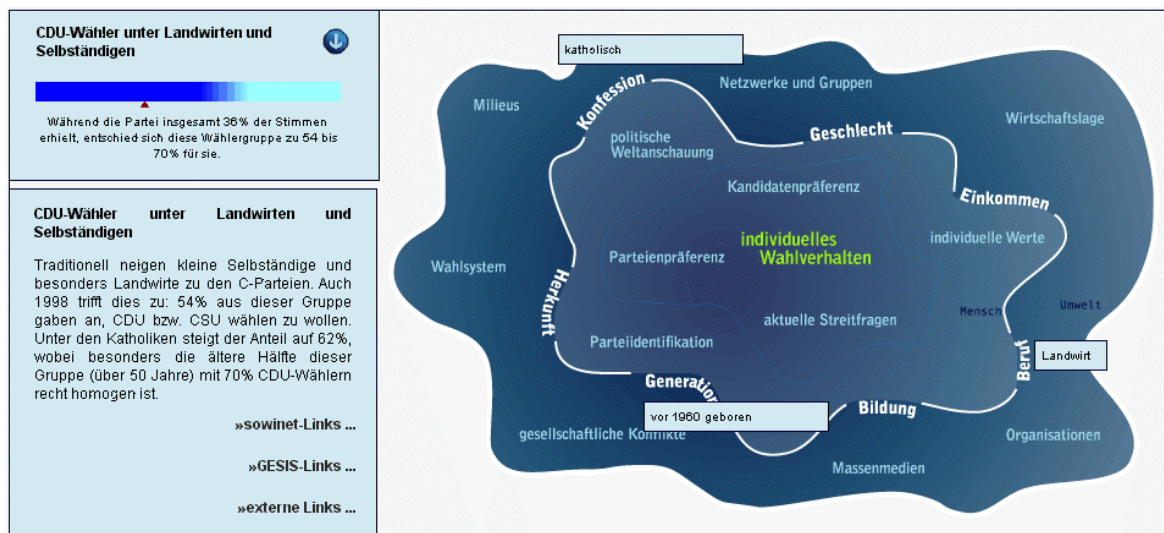


Abb. 4: Wählergruppenanalyse

5 Technische Aspekte der Umsetzung

5.1 Web-Realisierung

Die Realisierung des Wahlmodells besteht aus einer Kombination aus HTML, XML und Javascript. Die eigentlichen Inhalte des Modells sind dabei in einer XML-Datei abgelegt, die aus drei grundlegenden Elementen besteht:

1. Die eigentlichen Elemente des Modells wie z.B. „Kandidatenpräferenz“, „individuelle Werte“ oder „Wirtschaftslage“. Sie enthalten auch die zugeordneten Links.
2. Die Relationen zwischen den einzelnen Elementen.
3. Die Ausprägungen der Wählergruppenanalyse.

Diese Elemente werden aus einer externen XML-Datei über Javascript eingelesen und in den verschiedenen <div>-Bereichen der HTML-Seite entweder direkt als Text oder als Graphik umgewandelt präsentiert. Dabei erfolgen Zuordnungen wie die Kategorisierung der Links zu „sowinet“, „GESIS“ und „extern“ automatisch; ebenso wie die Erstellung des Prozentbalkens der Wählergruppenanalyse.

Der Vorteil dieser Vorgehensweise liegt in der Unabhängigkeit der Inhalte von der Darstellung. Dieser Vorteil ist in Zusammenhang mit XML in der Literatur inzwischen ebenso vielbeschrieben wie akzeptiert, weshalb hier nicht nochmals eine Lanze für ihn gebrochen zu werden braucht. Der Nachteil dieser Umsetzung soll nicht verschwiegen sein: Aufgrund der XML-Javascript-Kombination kann das Wahlmodell nicht mehr browserunabhängig präsentiert werden. Zum Einlesen einer externen XML-Datei wird ein Browser benötigt, der DOM-Javascript nutzt. Dies ist beim Microsoft Internet Explorer erst ab Version 5.5 und beim Netscape Navigator erst ab Version 6.2, und auch nur mit Hilfe von Tricks in der Programmierung, der Fall.

5.2 Statistischer Hintergrund der Wählergruppen-Analyse

Das Modell ist offen für ganz unterschiedliche Arten der Datenpräsentation. In der wissenschaftlichen Wahlforschung zielen empirische Analysen üblicherweise auf formale Hypothesentests und die Quantifizierung von Zusammenhangsstärken. Diese könnten durchaus unter dem vorliegenden Ansatz visualisiert werden, jedoch setzt die Interpretation einige methodische Kenntnisse voraus. Der hier gewählte Ansatz stellt dagegen auf die Bildung von anschaulichen Wählertypen ab, eine Quantifizierung erfolgt nur über Prozentsätze. Dabei wurden allein soziodemografische Merkmale zur Typenbildung herangezogen, unter anderem deshalb, weil diese auch im Alltag geläufig sind. Die Nähe zu gewissen Vorurteilen ist dabei kein Nachteil: Ob es stimmt, dass Bauern „schwarz“ oder Arbeiter „rot“ wählen, kann im Modell durch einen Klick auf die entsprechenden Wählergruppen leicht geprüft werden. Diese kognitiven Brücken führen durch die inhaltliche Differenzierung in den Begleittexten und durch die Einbindung in den theoretischen Rahmen schnell zu Vertiefungsmöglichkeiten.

Die Typenbildung beruht technisch-statistisch auf einem automatischen Segmentierungsverfahren.¹ Das Verfahren untersucht eine Anzahl von Variablen - hier die verschiedenen soziodemografischen Merkmale - auf ihren Zusammenhang mit einer „abhängigen“ Variablen, hier der Wahlabsicht. Dabei werden die Ausprägungen der verschiedenen Merkmale so zusammengefasst, dass die entstehenden Merkmalskombinationen, wie z.B. katholische Landwirte über 40 Jahren einerseits und protestantische Arbeiter unter 30 Jahren andererseits, sich hinsichtlich der abhängigen Variable Wahlverhalten möglichst stark voneinander unterscheiden. In der Regel ergeben sich so theoretisch plausible Merkmalszusammenfassungen, die hier direkt in die oben beschriebenen Wählertypen übersetzt werden können. Für diese Typen lassen sich nun die im Modell gezeigten Stimmenanteile der einzelnen Parteien direkt aus dem Programoutput ablesen.

Der weitgehend automatische Ablauf der Datenanalyse ist im Zusammenhang unseres Wahlmodells ein wesentlicher Vorteil. Dadurch wird eine relativ einfache und schnelle Wiederholung der Analysen mit anderen Daten möglich. So können ohne großen Aufwand für die zuliefernden Wissenschaftler etwa Aktualisierungen nach neuen Umfragen oder Vergleiche über Umfragedaten aus mehreren Jahren umgesetzt werden. Beides sind gerade in der Wahlforschung laufend wiederkehrende Anforderungen, die deshalb auch in populärwissenschaftliche Umsetzungen eingehen sollten.

6 Fazit

Das in diesem Artikel beschriebene Wahlmodell integriert die bei Informationsvisualisierung üblicherweise unabhängig voneinander behandelten Aspekte der Wissensvermittlung, Navigation und Analyse in einer Darstellung. Dabei bietet es einen weichen Übergang von einer relativ einfachen erklärenden Darstellung hin zu einem deutlich komplexeren Analysewerkzeug. Durch diese mehrschichtige Konzeption wird der Anwender einer Art explorativen Sog ausgesetzt, der ihn tiefer in die Materie hineinzieht.

Zum momentanen Stand ist das Wahlmodell integrierter Bestandteil von „sowinet.de“. Es sind zwei weitere Arbeitsschwerpunkte in Bezug auf das Wahlmodell geplant: Erstens sollen die Inhalte, hier v.a. die Linksammlung,

¹ Diese Prozedur namens CHAID (chi squared automated interaction detection, hier in der Umsetzung in SPSS CHAID 6.0.2. 1995) wurde auf Umfragedaten der „Politbarometer“ der Forschungsgruppe Wahlen aus den vier Wochen vor der Bundestagswahl 1998 angewendet (ZA o. J.).

noch weiter ausgebaut werden. Zweitens wird das Wahlmodell aus „sowinet.de“ entkoppelt, um es auch nach der Projektlaufzeit im Internet zur Verfügung zu stellen, beispielsweise als Unterrichtswerkzeug für die Sozialwissenschaften.

7 Literatur

- Bürdek, Bernhard E.; Dasan, Fatma; Denzinger, Jochen; Eibl, Maximilian; Jagodzinski, Wolfgang; Krause, Jürgen; Quandt, Markus (2001). sowinet.de – Interdisziplinäre Kooperation zur Erstellung einer breitenwirksamen Internet-Site zu Sozialwissenschaftlichen Themen. In: 7th Annual Meeting of the IuK Initiative Information and Communication of the Learned Societies <http://www.zpid.de/iuk2001/program/abstracts/Eibl.html> [1.8.2001].
- Bürklin, Wilhelm; Klein, Markus (1998). Wahlen und Wählerverhalten. Eine Einführung. Opladen.
- Däßler, Rolf; Palm, Hartmut (1998). Virtuelle Informationsräume mit VRML. Information recherchieren und präsentieren in 3D. Heidelberg.
- Eibl, Maximilian (2000). Visualisierung im Document Retrieval – Theoretische und praktische Zusammenführung von Softwareergonomie und Graphik Design. IZ-Forschungsberichte Bd. 3, Bonn.
- Eibl, Maximilian (2001). Interaction through Multimedia. In: Abridged Proceedings of the 9th HCI International Poster Sessions, S. 130-132.
- Eibl, Maximilian (2002). DEViD: A Media Design and Software Ergonomics Integrating Visualization for Document Retrieval. Erscheint in: Information Visualization, Bd.1, Nr. 2.
- Eibl, Maximilian, Mandl, Thomas (2002a). Niedrigdimensionale Dokumentkarten: Ein prototypisches Werkzeug zur flexiblen Visualisierung komplexer Datenbestände. Erscheint in: Proceedings des Internationalen Symposium für Informationswissenschaft 2002, Demo-Session.
- Eibl, Maximilian; Mandl, Thomas (2002b). Including User Strategies in the Evaluation of Interfaces for Browsing Documents. In: Journal of WSCG, Bd. 1, Nr. 2, S. 163-169.
- Elzer, P.; Krohn U. (1999). Visualization of Scientific Information in a Virtual Information Space. In: Proceedings of the HCI International'99, S. 122-126.
- Hearst, Marti; Petersen, J. (1996). Visualizing information retrieval results: a demonstration of the TileBar interface. In: Proceedings of the CHI '96, S. 394-395.
- Hemmje, Matthias; Kunkel, Clemens; Willett, Alexander (1994). LyberWorld - A Visualization User Interface Supporting Fulltext Retrieval. In: Proceedings of the ACM SIGIR'94, S. 249-259.
- Klein, Peter; Limbach, Tobias; Müller, Frank; Reiterer, Harald; Eibl, Maximilian (2002). Visualization of Metadata using the SuperTable + Scatterplot. Erscheint in: Proceedings des Internationalen Symposiums für Informationswissenschaft ISI 2002.
- Korfhage, R.R. (1991). To See or not to See - is That the Query? In: Proceedings of the ACM SIGIR'91, S. 134-141.

- Odgen W, Davis M, Rice S. Document thumbnail visualizations for rapid relevance judgments: When do they pay off? In: Proceedings of the 7th Text Retrieval Conference (TREC 7), S. 599-612.
- Quandt, Markus (im Erscheinen). sowinet.de - ein Internetjournal für die Sozialwissenschaften. Ziele, Konzept und experimentelle Umsetzung einer neuen Plattform. Erscheint in: Proceedings der ISKO 2001.
- Michard A. (1982). Graphical Presentation of Boolean Expressions in a Database Query Language: Design Notes and an Ergonomic Evaluation. In: Behaviour and Information Technology Nr. 3, S. 279-288.
- Reiners, Christine S.; Saborowski, Jörg (1997). Computervisualisierung, Raumvorstellung und Modelldenken im Chemieunterricht. In: Der Kanzler der Universität zu Köln (Hrsg.): Forschungsbericht Multimedia, Köln.
- Reiterer H, Mann TM, Mußler G. (2001). Visual Information retrieval for the WWW. In: Proceedings of the 9th HCI International, Bd. 1, S. 1150-1154.
- Roppel, Stephan (1996). Visualisierung und Adaption: Techniken zur Verbesserung der Interaktion mit hierarchisch strukturierter Information. Schriften zur Informationswissenschaft Bd. 33, Konstanz.
- Schnell, Rainer (1994). Graphisch gestützte Datenanalyse. München u. Wien.
- Spoerri A. (1994). InfoCrystal: Integrating Exact and Partial Matching Approaches through Visualization. In: Proceedings of the RIAO'94, S. 687-696.
- SPSS (1995). CHAID, Version 6.0.2. Chicago.
- Tufte, Edward (1983). The Visual Display of Quantitative Information. Cheshire, Connecticut.
- Wolff, Christian (1996). Graphisches Faktenretrieval mit Liniendiagrammen: Gestaltung und Evaluierung eines experimentellen Rechercheverfahrens auf Grundlage kognitiver Theorien der Graphenwahrnehmung. Schriften zur Informationswissenschaft Bd.24, Konstanz.
- ZA (o. J.). Zentralarchiv für empirische Sozialforschung: Politbarometer 1998, ZA-Studiennr. 3160. Köln.



Visualization of Metadata using the SuperTable + Scatterplot

*Tobias Limbach¹, Frank Müller¹, Peter Klein¹,
Harald Reiterer¹, Maximilian Eibl²*

¹University of Constance
Box D73,
78457 Konstanz, Germany
{Tobias.Limbach, Frank Müller,
Peter.Klein, Harald.Reiterer}@
uni-konstanz.de

²IZ Social Science Information Centre,
Bonn
eibl@berlin.iz-soz.de

Abstract

In this paper we will present a new approach combining different visualizations into a so called SuperTable with a Scatterplot. The goal was to improve user experience during the information seeking process. Originally INSYDER was a system to supply small and medium size enterprises with business information from the Internet by using a local meta search engine. The new prototype will be much more flexible concerning the operational area as well as the used visualizations. Instead of offering multiple visualizations to choose from INSYDER combines them into one, reducing the confusion of different visualizations.

1 Introduction

Some of the main challenges of the Web are problems related to the user and their interaction with the retrieval system. Conventional document retrieval systems return long lists of ranked documents that users are forced to sift through to find relevant documents. The majority of today's web search engines follow this paradigm. The notoriously low precision of web search engines coupled with the ranked list presentation makes it hard for users to find the information they are looking for. There are basically two problems: *how to specify a query and how to interpret the answer provided by the system*. Surveys have shown that users face problems with the current paradigm of



information retrieval systems for Web search simply presenting a long list of results. These long lists of results are not very intuitive for finding the most relevant documents in the result set. These empirical findings motivated us to develop a new type of user interface for Web retrieval that supports the user in the information seeking process by providing special visualizations in addition to the traditional result list. Systems combining the functionality of retrieval systems with the possibilities of information visualization systems are called *visual information retrieval systems*.

This paper presents our main design ideas developing a visual information retrieval system called INSYDER [Reiterer et al. 2000]. Chapter 2 summarizes our experiences developing the INSYDER system and presents the major visualizations. Based on a comprehensive empirical user test with 40 users we have made a redesign of the original INSYDER system. One of the main results was that a combined visualization based on the result table would fulfill most of the users needs to filter, group and select the results. Chapter 3 presents in detail the redesigned system, which uses a combination of the original visualizations, the so called "SuperTable". We are now developing two different versions of the "SuperTable" using different strategies showing the granularity of details (levels of details versus a smooth change of granularity approach). We also provide an overview about related Usability Engineering activities. Chapter 4 shows related work that has influenced our work. Conclusions and outlook are given in chapter 5 where also a system adapted to mobile computing is mentioned.

2 INSYDER - A Visual Information Retrieval System for the Web

The main goal of the INSYDER¹ project is to create a solution to supply small- and medium-size enterprises with business information from the Web. To make the information accessible, the basic idea behind INSYDER is a software-plus-content approach. The software is a local meta-search engine with functions for searching and crawling HTML- and TXT-based information, monitoring changes of retrieved documents, handling news and bookmarks, and last but not least managing all this in a topic-oriented way in Spheres Of Interest (SOIs). "Content" means country- and industry-branch-specific predefined SOIs with selected bookmarks, collections of starting

¹ The INSYDER project was funded by the European Commission under the Fourth Framework of the Esprit Program, Domain 1, Task 1.9 Emerging Software Technologies. Project No. 29232.

points like search engines and URL-lists, specific thesauri to improve the relevance ranking of the semantic analysis module, or rule files to classify hits by user definable host-types. On the whole INSYDER is created as a country- and industry-branch-specific adaptable system to find, to evaluate, to filter, to manage, and to monitor relevant business information from the Web.

The final implementation of the INSYDER system [Reiterer, Mußler, Mann 2001] included five components for the presentation of search results: a HTML-List, a ResultTable, a Scatterplot, a BarGraph, and a SegmentView with two modes: TileBars and StackedColumn. For details-on-demand functions there are also a segment tooltip, a document tooltip, a text window, and a browser. During the development of the INSYDER system it was not intended to come up with new visual metaphors supporting the retrieval process. We tried to select expressive visualizations keeping in mind the target users (business analysts), their typical tasks (to find business data in the Web), their technical environment (typically a desktop PC and not a high-end workstation for sophisticated graphic representations), the type of data to be visualized (document sets and text documents), and minimal necessary training. The major challenge from our point of view was to combine in a smart way the selected visualization supporting different views on the retrieved document set and the documents themselves. The primary intention was to present additional information (metadata) about the retrieved documents to the user in a way that is intuitive, may be quickly interpreted, and can scale to large document sets. We have used two different approaches depending on the additional information presented to the user:

- Predefined document attributes: E.g. title, URL, server type, size, document type, date, language, relevance. The primary visual structures to show the predefined documents attributes are the Scatterplot and the Result Table.
- Query terms` distribution: This shows how the retrieved documents related to each of the terms are used in the query. The primary visual structures to show the query terms` distribution are the BarGraph, the TileBar and the Stacked Column.

Another important difference of our INSYDER system compared to existing retrieval systems for the Web was the comprehensive visual support of different steps of the information seeking process. The visual views used in INSYDER support the user's interaction with the system during the formulation of the query. For example Figure 1 shows the visualisation of related terms of the query terms with the help of a graph (Mußler 2002).

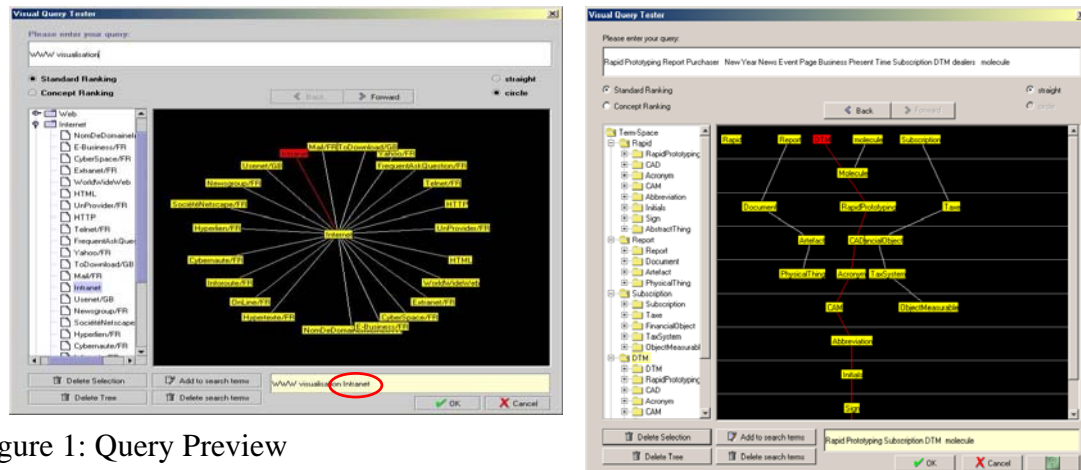


Figure 1: Query Preview

Further views are shown in Figure 2 and could be used during the review of the search results: visualisation of different document attributes like date, size, relevance of the document set with a ResultTable (Figure 2, top left), a Scatterplot (Figure 2, bottom left), Bar Graphs (Figure 2, top right) or a visualisation of the distribution of the relevance of the query terms inside a document with TileBars (Figure 2, bottom right), and during the refinement of the query (e.g. visualisation of new query terms based on a relevance feed-back inside the graph representing the query terms).

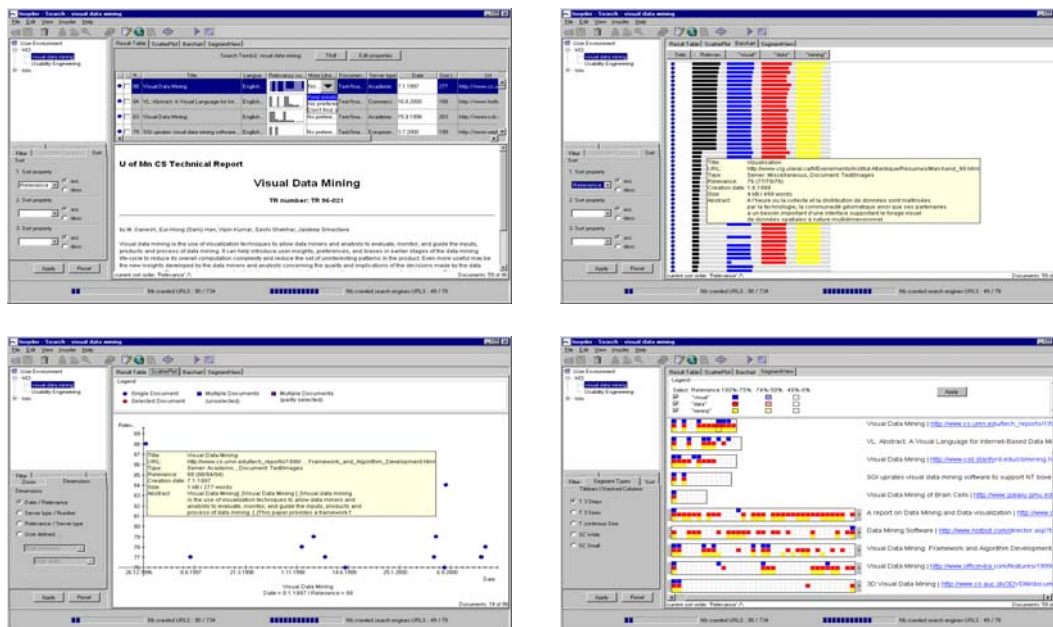


Figure 2: Views of INSYDER

The visual information seeking system INSYDER is not a general-purpose system like traditional search engines (e.g. AltaVista). Like mentioned before, its context of use is to support small and medium sized enterprises (SMEs) of specific application domains finding business information on the Web. With

the findings of general empirical studies (Nielsen 1997), (Pollock, Hockley 1997) the results from a field study, which was conducted at the beginning of the project, using a questionnaire that has been answered by 73 selected companies (SMEs) in Italy, France and Great Britain, our aim was to understand the context of use (ISO 9241 Part 11) following a human-centred design approach (ISO 13407).

The primary goal of a summative evaluation with 40 users was to determine the usability of the visualization concepts in dependency of different factors. A second goal was to identify problems with the visualization ideas and components used in the INSYDER system, and to collect suggestions for improvements. The usability evaluation part of the study was focused on the added value of the visualizations (Scatterplot, BarGraph, TileBar, Stacked-Column) in terms of their effectiveness (accuracy and completeness with which users achieve task goals), efficiency (the task time users expended to achieve task goals), and subjective satisfaction (positive attitudes to the use of the visualization) for reviewing Web search results.

The results from the evaluation of the INSYDER System [Mann 2002] pointed out some difficulties of user interaction with the system, e. g.: More than 50% of the users voted for the ResultTable, when asked, which visualization performed best. Other visualizations were helpful as an addition to the ResultTable, but not as primary tools. When studying the expected value of a component, it can be concluded that in the Visualization plus ResultTable conditions, where the user had the possibility to decide which component to use, in the majority of cases both components were used. When analyzing usage times under these conditions, the ResultTable was the favorite component of the users. It was used under all three user interface conditions with Scatterplot, BarGraph, and SegmentView for more than 50% of the overall task time. Interpreting usage time as an indicator of expected value, the expected value of the ResultTable seemed to be higher than that of the other components for the users. Switching between completely different visualizations confused the users. So we tried to find a possibility to combine the regular table view with other views like the BarGraph or the SegmentView.

3 Redesign of INSYDER - The SuperTable + Scatterplot

Based on the empirical findings of the evaluation the following improvements of the overall system will be discussed (see also [Klein et al. 2002]):

- The Integration of Table Lens, ResultTable, HTML-List, BarGraph, and SegmentView into one component called „SuperTable”.
- Improvements for the Scatterplot component.

The users requested a number of features for the BarGraph and the Segment-View already implemented in the ResultTable. All this could also be implemented in the ResultTable. So the proposed “SuperTable” must integrate a multiple focus Table Lens, the HTML-List, a BarGraph and the SegmentView with TileBars and StackedColumn in a way that allows easy manipulation of the table. There were to be different predefined zoom-levels. Starting for example with a BarGraph view in a first zoom level, revealing document details in a one-line-per-document-mode for the whole table in a second level, a several-line-per-document-mode in a third level, and document information revealing segment details in a fourth level. In addition, tooltips or lens mechanisms allow easy inspection of details for single documents. Power users can use multiple focus possibilities for comparisons. The SuperTable as an integrated component will be more complex than the former single components. With adequate design a person who is starting with the system may not be aware of this complexity, and therefore will be distracted. The integration will minimize context-switching effort and is able to allow a smooth learning curve from beginner mode to power user mode.

Nevertheless, the scatterplot may also be improved by packing more information into the display, for example, by using partially filled circles in the Scatterplot, as done by [Fishkin, Stone 1995] with their real-valued queries. If this is done, the overall relevance of a document will always be visible, even when the dimension is not mapped to one of the axes. With the callout lens from the same authors, an exploder for document groups can be integrated allowing to fulfill one of the requests formulated by the users. The callout lens can be used to explode “clumps” of icons or document group symbols. With this and other enhancements, the Scatterplot could give the user a quick overview of all search results, and offering the user a good variety of controls (e.g. defining personal views, zooming, selecting) to reduce the amount of hits to a smaller group of interesting documents. These documents can then be selected by the mouse and analyzed in more detail in the SuperTable. An enhanced Scatterplot with additional lens mechanisms tightly coupled with a SuperTable and a document browser are our main redesign ideas.

Visualising the Level Concept: Discrete Change of Modalities

Based on the empirical findings of the evaluation we have decided to integrate the ResultTable, BarGraph, and SegmentView into one visualization called „SuperTable” and to improve the Scatterplot.

In the user test the users requested a number of features for the BarGraph and the SegmentView already implemented in the ResultTable. All this could also be implemented in the ResultTable. Therefore, the proposed SuperTable integrates the concept of a distortion-based table, the BarGraph and the SegmentView (with TileBars and StackedColumns) in a way that allows easy manipulation of the table.

The redesign of the INSYDER visualizations combines the SuperTable + Scatterplot into one single window offering different brushing techniques between them. Therefore, the Scatterplot will supplement the SuperTable by giving the user a quick overview of all search results, and offering the user a variety of controls (e.g. defining own views, zooming, selecting, filtering) to reduce the amount of hits to a smaller group of interesting documents. These documents can then be selected by the user and analyzed in more detail in the SuperTable.

The enhanced Scatterplot with additional lens mechanisms (e.g., a magic lens for filtering operations [Fishkin, Stone 1995]), distortion techniques in both X- and Y-dimensions (allowing a smooth transition between focus and context) and the radial MDP visualization tightly coupled with the SuperTable and a document browser (showing the detailed document with keyword highlighting) are our main redesign ideas.

The first design variation of the SuperTable with the level concept, which is combined with the Scatterplot and a Browser window on a single panel, is shown in Figure 3.

At first sight, the SuperTable doesn't differ from the original ResultTable: Every row represents the metadata of a web search result – in that case a HTML-page. The columns describe all characteristics concerning this document, e.g. title, URL, date, size and so on. The main feature of the initial view is that no text at all will be seen in the table. There are just a few multicolored bars representing numeric and textual data.

The difference between the origin ResultTable and the SuperTable consists of four different views representing the "level of detail" for documents. We call

this method of looking closer and closer at the details of a document the "Focus of Interest". The more you want to know about a document, the deeper you have to look, i.e. the higher the level has to be. It is possible to focus single documents, several documents or the whole list of documents.

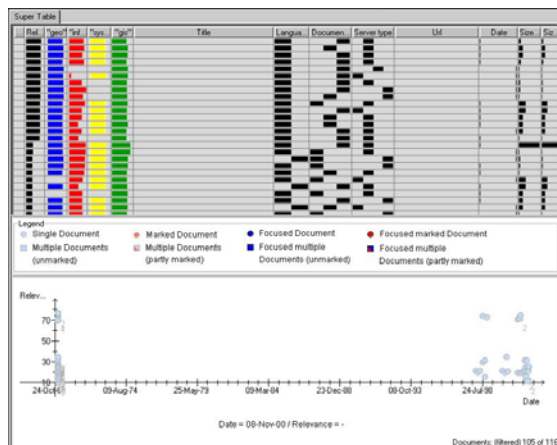


Figure 3: SuperTable Level 1

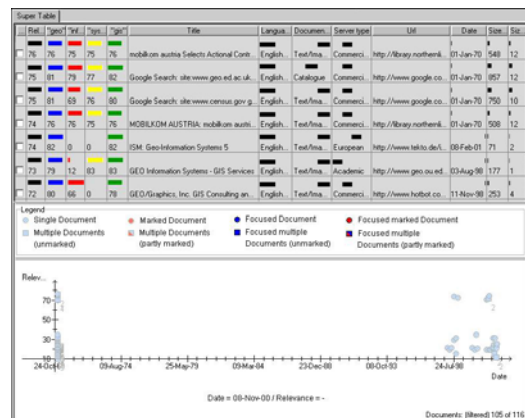


Figure 4: SuperTable Level 2

Level 1 represents an overview over all documents (Figure 3). All rows are as small as possible, so that in the best case all documents fit on the available space. Corresponding to the number of documents, the height of the rows can vary. Usually the rows will be too small to hold text, so only bars will be displayed. The length and the position of the bars encode various characteristics of the document depending on the type of data they represent. The length of the bars (representing numeric data like size, relevance) is equivalent to their numeric values. Nominal attributes can also be represented by a bar. For a few attributes we can code their values through position; for example, the language which may be English or French in our application (the left half of the cell means "English" and the right one means "French"). If there were too many different nominal values, visualization would be too confusing which is for example the case with the title. Therefore, this column will be empty. Different colors can additionally encode different search terms.

On Level 2 (Figure 4) more information will be visible in form of text completing the visual representation of the multicolored bars. Now numeric values add detailed information about the bar displays from the initial, graphics-only display. Title and URL are now readable, but only up to the width of the respective column. All wider texts become truncated, marked by three dots.

Level 3 provides the opportunity to read the whole text of those characteristics, which had to be abbreviated because of their size (Figure 5). Visualizations were cut off to gain space, so that title, URL and abstract are now completely visible. In addition a new column is introduced, the so-called "Rele-

vance Curve". It represents a two-dimensional chart of the whole document by dividing the document into a number of segments, e.g. sentences, subordinate clauses, etc. The height of single bars encodes the overall relevance for each individual segment.

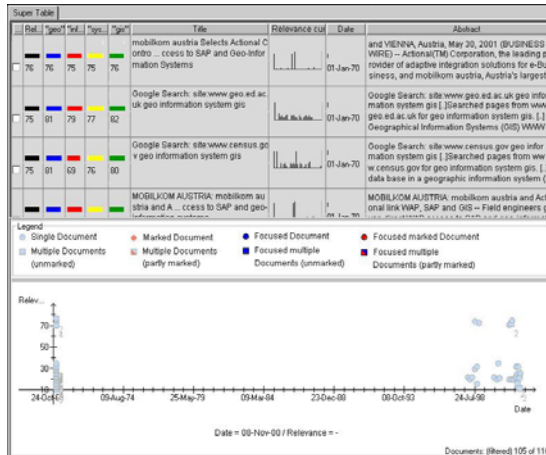


Figure 5: SuperTable Level 3

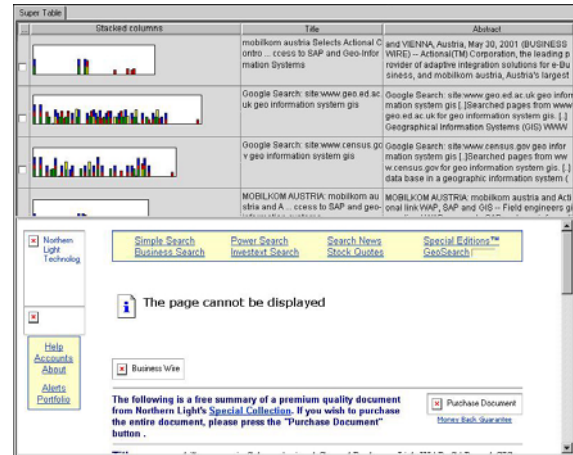


Figure 6: SuperTable Level 4

Level 4 (Figure 6) displays only the most important values of a document. The title, an abstract and an extension of the relevance curve, the so-called Segment View which uses stacked columns or TileBars [Hearst 1995]. Which kind of query term distribution visualization will be used depends on the user preferences. In this case, every segment has the same length, not varying from one document to another. Consequentially the length of the SegmentView differs from document to document, always corresponding to the real length of the document. The stacked columns or TileBars are colored according to the colored search terms in the former levels. So it is easier to discover the segments where all search terms can be found, not just a few of them. It is easy to spot segments which include only one or two search terms and discern these from parts including all terms. Some terms may be seen more important than others, so a segment with a lower relevance can be important although not all terms are included. Per default, instead of the Scatterplot a preview browser is offered showing the selected document or per default the first document of the SuperTable. This offers the user the greatest level of detail showing him the text of the document (without images) enhanced by keyword highlighting.

Smooth Change of Modalities: the Granularity Concept

Visualizations in document retrieval have one common flaw: the change of modalities during a search. The query is formulated in a textual way, for example by a simple list of search terms, a Boolean combination of search terms, a form, a SQL-statement and so on. Only a few systems allow a mere graphical input. In a next step a graphical version of the result set is presented.

Here the user can see the effects of his query on the result set. Some visualization even allows manipulating the result set. Finally, the documents themselves are presented textually, usually in a selectable list of titles.

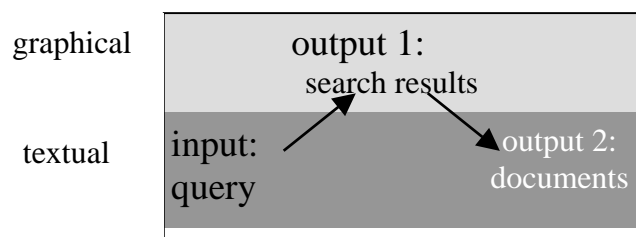


Figure 7: Change of modalities during a search

Thus, the user has to switch at least twice between modalities. The first time in order to get from his query formulation to the visualization and the second time in order to get from the visualization to the actual documents. Nevertheless we estimate that during a normal search, he will use iterative retrieval and therefore he will have to switch between the modalities much more often. Each time he has to connect the textual representation to the graphical mentally.

This cognitive load of connecting the modalities raises the question whether in this context visualization does make sense at all. In general, visualizations are employed because of the human visual capacities [Card, Mackinley, Shneiderman, 1999]. The human mind is able to cope with a comparably high amount of visual data and has serious problems of juggling with the same amount of textual data. Based on this fact, visualizations in document retrieval are mostly employed in order to explain the search result. Patterns and exceptions in the result set can easily be detected and the query reformulated appropriately.

However, this advantage of visualization is won at the cognitive expense of the transfer between the modalities. Here we want to introduce a possibility of reducing this expense. Systems like InfoCrystal [Spoerri 1994] and DEViD [Eibl 2000] integrated the search into the visualization and thus reduced the problems of the first transfer. INSYDER might be evolved to solve the problem of the second transfer from the visualization to the textual output.

The level concept presented above tries to combine the visualizations employed by INSYDER in order to ease their interpretation. In a design variant, this level concept is further smoothened by introducing the concept of granularity. Granularity is a term used in photography to describe the accuracy of

lows to show the abstract with keyword highlighting of the query terms. In the TileBar the user may jump to the relevant parts of the text he wants to see.

Step 6 (Figure 13): In the last step only two columns are left: the visualization column and the text column are merged with another. The maximal granularity (the whole text of the document) is reached. To retain the advantage of the visualization, the first two columns of step 5 are combined in a single one offering as much space as possible for the final text presentation, with the query terms highlighted.

Two goals can be achieved with this design variant of the SuperTable: First, the several visualizations do not appear as distinct anymore. On the contrary, the user has the feeling that he is working with one single visualization in different states.

Secondly, the SuperTable solves a problem, which seemed to be immanent to visualizations in document retrieval: the change of modalities. During the retrieval process the user has to get past a modality change twice: one time from the textual input of the query to the visualization and a second time from the visualization to the textual output of the documents. These changes are usually rather harsh, because researchers focused too much on the visualizations themselves leaving aside their embedding in the retrieval process.

Past and Upcoming Usability Tests

Early in the redesign of the original INSyder project the decision for an early usability study was taken. At this stage, there was a general idea about the design and further concepts, and keeping the results of the latest usability study in mind the motivation to rethink basic functions of INSyder.

With the new visual approach described in Chapter 3.1 (SuperTable + Scatterplot) a usability study was carried out during a week in August 2001. The participants (n=8) were students and co-workers of the department of computer and information science at the University of Constance. The study material consisted of several laminated paper-mockups, displaying different stages of the visualization and corresponding several elements for interactivity (e.g. printout of context menu, printout of tool-tips etc. Basically the material presented throughout Figures 3 - 6). After a brief introduction by the moderator the participants went through different tasks, describing their action and thoughts by thinking-aloud. When the “system” (the prototype) had to respond to user actions, the GUI elements were put into the mockup. After the tasks were completed, the participants were asked to fill out a questionnaire

and were briefly interviewed to their overall satisfaction, suggestions and remaining questions. Everything was documented by minute taking and a video camera.

Seven out of eight participants stated that they could imagine to work with a system like the prototype presented. A third voted the visualizations as a *very helpful* support to search tasks, the rest as *helpful* support. The most helpful level for the participants was Level 3 (see Fig. 5), the least helpful was Level 1. Throughout the study, the Scatterplot, or better, the combination of SuperTable and Scatterplot was regarded as helpful or even very helpful. The change of detail level made surprisingly less problems than expected. With these important results in mind the decision to build a Html-based Prototype with more interactive elements and the look-and-feel of a real Java-Application was taken. This should give a chance to explore and research more about familiar and unfamiliar concepts.

Current work involves two Html-based demonstrators, one for the SuperTable plus Scatterplot, one for the Granularity Concept plus Scatterplot and a prototypical realization in Java. The demonstrators support the same basic functionality and are considered to be comparable. Both have and are currently undergoing minor design changes, mostly influenced by the Scenario Based HCI approach (Rosson, Carroll 2002). The Java working prototype will work as a technical “proof-of-concept” realization and also as a basic version for further metadata visualization frameworks. The next step in usability testing will be a competitive ease-of-use analysis between the two demonstrators. Here we expect the users to be challenged by the granularity concept, but keeping the ease of acceptance of the change of detail level approach in mind, are confident that users can successfully work with it.

4 Related Work

The original INSYDER system has been influenced by different existing systems. The use of the *Scatterplot* was mainly inspired by the visual information seeking systems Envision [Nowell, France, Hix et al. 1996] and FilmFinder [Ahlberg, Shneiderman 1994]. The use of the *BarChart* was mainly inspired by the work of [Hearst 1995].

The SuperTable idea has been influenced by multifocal approaches using focus-plus-context techniques as can be found in a number of tabular data representations. Interactive tables like the Table Lens [Rao, Card 1994] or FOCUS [Spence, Beilken, Berlage 1996] are typical examples. Textual and graphical

representations of the data are used in both systems. Focus-plus-context allows showing more cells of the data table on the screen than without this technique. In both systems, the coherence of rows and columns and their labels is preserved when distorting parts of the view. The graphical elements are used for pattern recognition when working with quantitative variables. Whereas in the Table Lens the cases are displayed in rows, in FOCUS they are displayed as columns.

5 Conclusion and Outlook

The advantage of INSYDER was the variety of visualizations used to support users in their web search. Our redesign combines these visualizations with the widely spread list-like layout in the SuperTable. New possibilities are given to find the most appropriate document for the current task in an environment users are accustomed to. The evaluation of the redesign proved that we're on the right way. Both approaches presented in this paper give us the possibility to improve the advantages of this combination in a more detailed way. Additionally a highly sophisticated data model will enable us to adapt INSYDER to a wide range of fields like stock market, medical data mining or geographic information systems. Concerning the last we are currently adapting our visualization framework to a new application domain in a project called INVISIP (Information Visualization in Site Planning)³. The aim of this project is to support users in the retrieval process of complex application areas such as site planning and to facilitate graphic-interactive access to geodata archives. Here, the idea is to provide information visualization techniques in the different information retrieval phases to locate appropriate geodata, which is necessary to solve planning tasks, e.g. generation of ecological, environmental or socio-demographic reports. The SuperTable + Scatterplot will be introduced in a 3D GeoLibrary [Göbel, Haist, Goebel 2002][Göbel, Haist, Reiterer, Müller 2002] as one new information visualization techniques to support users during the different information retrieval phases, especially search result presentation and comparison of search results (hints for query modification). Another field of work will be a minimalistic Java version, which will be implemented for and tested in NTT's DoCoMo DoJa 2.0 API Development Kit for DoCoMo's 504i mobile telephones. Especially in the domain of information appliances the "traditional" result list approach seems cognitively demanding and unsatisfactory. The granular SuperTable in a reduced version seems promising

³ The project is funded by the European Commission under the Fifth Framework of the IST Program, Project No. IST-2000-29640, www.invisip.de.

enough to start some serious thoughts about early prototyping and evaluation of possible design scenarios.

After the redesign of both demonstrators is finished we will perform an evaluation to compare the two redesign ideas and to decide which one will have the most benefit for the users. New visualizations will be added to expand the variety of visual aids. Furthermore we plan to support visual query generation and more graphics based filter and sort mechanisms to improve the ease of use.

6 References

- [Ahlberg, Shneiderman 1994] Ahlberg, Christopher; Shneiderman, Ben: Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays. In: Adelson, B.; Dumais, S.; Olson, J. S. (Eds.): CHI 1994: Conference Proceedings Human Factors in Computing Systems. Conference: Boston, MA, April 24-28 1994. New York (ACM Press) 1994. p. 313-317.
- [Card, Mackinley, Shneiderman, 1999] Card, Stuart; Mackinley, Jock; Shneiderman, Ben (Eds.) (1999). Readings in Information Visualization. Using Vision to Think. San Francisco.
- [Eibl 2000] Eibl Maximilian: Visualisierung im Document Retrieval. Forschungsberichte Band 3, IZ Informationszentrum Sozialwissenschaften, Bonn, 2000
- [Fishkin, Stone 1995] Fishkin, Ken; Stone, Maureen C.: Enhanced Dynamic Queries via Movable Filters. In: Katz, Irvin R.; Mack, Robert L.; Marks, Linn et al. (Eds.): CHI 1995: Conference Proceedings Human Factors in Computing Systems. Conference: Denver, CO, May 7-11 1995. New York (ACM Press) 1995. p. 23-29.
- [Göbel, Haist, Goebel 2002] Göbel, Stefan; Haist, Jörg; Goebel, Christoph: GeoCrystal : Graphic-Interactive Access to Geodata Archives. Proceedings SPIE 2002 – Visualization and Data Analysis, San Jose, CA, 2002.
- [Göbel, Haist, Reiterer, Müller 2002] Göbel, Stefan; Haist, Jörg; Reiterer, Harald; Müller, Frank: INVISIP: Usage of Information Visualization Techniques to Access Geospatial Data Archives. Accepted Paper at the DEXA 2002 Conference, September 2-6, 2002, Aix en Provence, France
- [Hearst 1995] Hearst, Marti A.: TileBars: Visualization of Term Distribution Information in Full Text Information Access. In: Katz, Irvin R.; Mack, Robert L.; Marks, Linn et al. (Eds.): CHI 1995: Conference Proceedings Human Factors in Computing Systems. Conference: Denver, CO, May 7-11 1995. New York (ACM Press) 1995. p. 59-66.
- [Klein et al. 2002] Klein, Peter; Müller, Frank; Reiterer, Harald; Eibl, Maximilian: Visual Information Retrieval with the SuperTable + Scatterplot. In: 6th International Conference on Information Visualization (IV 02). London (IEEE Press) 2002, p.70-75
- [Mann 2002] Mann, Thomas M: Visualization of Search Results from the World Wide Web, PhD Thesis, University of Konstanz, 2002,
<http://www.ub.uni-konstanz.de/kops/volltexte/2002/751/>

- [Mußler 2002] Mußler, Gabriela: Retrieving Business Information from the WWW. PhD Thesis, University of Konstanz, 2002.
- [Nielsen 1997] Nielsen, Jacob: Search and You May Find.
<http://www.useit.com/alertbox/9707b.html> [1999-03-18].
- [Nowell, France, Hix et al. 1996] Nowell, Lucy T.; France, Robert K.; Hix, Deborah et al.: Visualizing Search Results: Some Alternatives to Query-Document Similarity. In: Frei, Hans-Peter; Harman, Donna K.; Schäuble, Peter et al. (Eds.): SIGIR 1996: Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. Conference: Zürich, Switzerland, August 18 -22 1996. New York (ACM Press) 1996. p. 67-75.
- [Pollock, Hockley 1997] Pollock, A.; Hockley, A.: What's Wrong with Internet Searching. D-Lib Magazine, 1997, <http://www.dlib.org/dlib/march97/bt/03pollock.html> [1999-02-01].
- [Rao, Card 1994] Rao, Ramana; Card, Stuart K.: The Table Lens. Merging graphical and symbolic representations in an interactive focus + context visualization for tabular information. In: Adelson, B.; Dumais, S.; Olson, J. S. (Eds.): CHI 1994: Conference Proceedings Human Factors in Computing Systems. Conference: Boston, MA, April 24-28 1994. New York (ACM Press) 1994. p. 318-322.
- [Reiterer et al. 2000] Reiterer Harald; Mußler Gabriela, Mann Thomas, Handschuh Siegfried: INSYDER - An Information Assistant for Business Intelligence, Proceedings of the 23 Annual International ACM SIGIR 2000 Conference on Research and Development in Information Retrieval, ACM press, 2000, pp.112-119
- [Reiterer, Mußler, Mann 2001] Reiterer, Harald; Mußler, Gabriela; Mann, Thomas M.: A visual information seeking system for Web search. In: Oberquelle, Horst; Oppermann, Reinhard; Krause, Jürgen (Eds.): Mensch & Computer 2001. Conference: Bad Honnef, Germany, March 5-8 2001. Stuttgart (Teubner) 2001. S.297-306.
- [Rosson, Carroll 2002] Rosson, Mary Beth; Carroll, John M.: Usability Engineering. Scenario-Based Development of Human-Computer Interaction. San Diego (Academic Press) 2002
- [Spenske, Beilken, Berlage 1996] Spenske, Michael; Beilken, Christian; Berlage, Thomas: FOCUS: The Interactive Table for Product Comparison and Selection. In: UIST 96: 9th ACM Symposium on User Interface Software and Technology New York (ACM Press) 1996. p. 41-50.
- [Spoerri 1994a] Spoerri, Anselm (1994a). InfoCrystal: Integrating Exact and Partial Matching Approaches through Visualization. In: RIAO'94 Proceedings, Oct.11-13, 1994 New York (NY), S.687-696.
- [Veerasamy, Navathe 1995] Veerasamy, Aravindan; Navathe, Shamkant B.: Querying, Navigating and Visualizing a Digital Library Catalog. In: Digital Libraries 1995: The Second Annual Conference on the Theory and Practice of Digital Libraries. Conference: Austin, TX, June 11-13 1995.
<http://www.csd.tamu.edu/DL95/papers/veerasamy/veerasamy.html> [1999-03-24]



Concept Extractor - Ein flexibler und domänenspezifischer Web Service zur Beschlagwortung von Texten

*Lukas C. Faulstich¹, Uwe Quasthoff², Fabian Schmidt¹,
Christian Wolff³*

¹pepper words GmbH,
Ndl. Leipzig
Karl-Heine-Str. 99
04229 Leipzig
{faulstich,fschmidt}@pepper-
words.de

²Universität Leipzig,
Institut für Informatik,
Augustusplatz 10/11,
04109 Leipzig
quasthoff@informatik.uni-
leipzig.de

³Universität Regensburg
Institut für Medien-, Informations-
und Kulturwissenschaften
93040 Regensburg
christian.wolff@sprachlit.uni-
regensburg.de

Zusammenfassung

Der Beitrag beschreibt ein flexibles und modulares System zur automatischen Beschlagwortung von Texten, das auf einer Text Mining-Engine aufbaut. Dabei liegt eine Methode der differentiellen Corpusanalyse zugrunde: Der zu verarbeitende Text wird im Vergleich mit einem umfangreichen Referenzcorpus analysiert und Unterschiede in relativen Häufigkeitsklassen dienen der Auswahl geeigneter Schlagworte. Zusätzlich kommen Datenbanken zum Einsatz, die eine Expansion von Termen hinsichtlich Grundform, Schreibvarianten, Synonymen und Mehrwortbegriffen erlauben. Das System ist als *web service* realisiert und lässt sich problemlos in Content Management-Systeme integrieren.

Abstract

We describe a flexible and modular system for keyword extraction and attribution which operates on top of a text mining engine. Texts are analysed in comparison with a large reference corpus and key words are determined using a frequency based method for determining relative term significance. Additionally, selected terms may be expanded using large knowledge bases on inflected forms, orthographic variants, synonyms and multi word terms. This solution is realised as a web-based service which can easily be integrated into existing content management systems.



1 Einleitung

Die Beschlagwortung redaktioneller Texte, die für verschiedene Medienkanäle aufbereitet und mittels *content syndication* auf unterschiedliche Weise verwertet werden (*cross media publishing*, cf. [Simon 02]), stellt hohe Anforderungen an eine dem jeweiligen Medium angemessene Beschlagwortung. Typische Szenarien umfassen dabei die mehrfache Verwertung medialer Inhalte (hier: redaktionelle Texte) über unterschiedliche Kanäle, z. B. durch

- Publikation auf Websites,
- Lieferung themenspezifischer Inhalte als push-Service per SMS,
- Bereitstellung von Texten für Videotextdienste oder
- die Verwendung in Printmedien.

Für die jeweilige Verwendung gelten jeweils auch unterschiedliche Anforderungen an die inhaltliche Aufbereitung der Texte im Sinne der Informationserschließung, sei es, dass Text für den internen redaktionellen Gebrauch in einem Content Management-System zu beschlagworten ist, sei es, dass eine Beschlagwortung zu erstellen ist, die eine gute Erschließung über Suchmaschinen zulässt (Verwendung von Schlagworten in Meta-Tags der im Web publizierten HTML-Seiten).

Im Folgenden wird ein Beschlagwortungsserver beschrieben, der auf einer mehrschichtigen Analyse von Textcorpora durch Text Mining-Verfahren aufbaut, eine flexible agentenbasierte Beschlagwortungslösung anbietet und als Web Service (cf. [Preece & Decker 02]) verfügbar und damit in unterschiedliche Content Management-Systeme integrierbar ist.

Dabei wird zunächst auf die corpuslinguistischen Grundlagen eingegangen (Kap. 2). Kap. 3 beschreibt Aufbau, Funktionsweise und technische Umsetzung des Beschlagwortungsservers und Kap. 4 gibt praktische Anwendungsbeispiele.

2 Corpuslinguistische Grundlagen

Die Verfügbarkeit elektronischer Texte hat im vergangenen Jahrzehnt zu einer Renaissance der Corpuslinguistik geführt, wie u. a. auch das schnelle Anwachsen internationaler Fachtagungen wie der *Conference on Language Resources and Evaluation* (LREC) zeigt. Mittlerweile liegen nicht nur für viele

Sprachen Referenzcorpora vor, auch die Bemühungen und die Standardisierung von Corpusaufbau und Analyse zeigen Erfolge (cf. [Atkins et al. 02]).

Für die Problematik der automatischen Beschlagwortung von Texten sind corpuslinguistische Verfahren deshalb von Interesse, da sie im Vergleich mit den bekannten Verfahren zur Textindexierung wie dem *vector space model* (cf. [Salton 83], [Baeza-Yates & Ribeiro-Neto 99]) Corpora als zusätzliche Bezugsgröße der Textanalyse einführen und, insofern Corpora nicht nur als Rohdatensammlung, sondern als strukturierter Informationsspeicher zur Verfügung stehen, auch für die Beschlagwortung relevante Informationen bereitstellen können (z. B. Grundformen, Synonyme, Sachgebietsangaben etc.).

2.1 Vergleich von Corpora

Für die dynamische Beschlagwortung von Texten ist dabei der Aspekt des *Vergleichs* unterschiedlicher Corpora von Bedeutung: Text Corpora lassen sich u. a. anhand Faktoren wie Umfang, Art, Anzahl oder Bezugszeitraum der in ihnen enthaltenen Dokumente beschreiben. Durch Vergleich unterschiedlicher Corpora lassen sich dabei Erkenntnisse über die Beschreibungsadäquatheit von Begriffen bezüglich einzelner Dokumente eines Corpus gewinnen.

Mit [Rayson & Garside 00:1] kann man zwei Typen des Corpusvergleichs unterscheiden:

- Vergleich zweier Corpora ähnlicher Größe, die sich hinsichtlich eines Parameters (z. B. Erhebungszeitraum) unterscheiden. Ein solcher Vergleich kann z. B. in der Semiometrie oder Trendforschung eingesetzt werden, um zu bestimmen, inwieweit sich Trends durch geänderten Sprachgebrauch (Wortverwendungshäufigkeiten, unterschiedlicher Vokabularaufbau) nachweisen lassen.
- Vergleich eines kleineren gegen ein größeres (normatives) Corpus, z. B. bei der Differenzierung zwischen Sprachgebrauch in einer Fachdomäne im Vergleich mit einem aus allgemeinsprachlichen Texten aufgebauten Corpus.

Für das nachfolgend beschriebene Beschlagwortungssystem ist der zweite Fall, d. h. der Vergleich von kleinerem Fachcorpus mit einem deutlich größeren normativen allgemeinsprachlichen Corpus der Ausgangspunkt. Als normative Corpusgrundlage dient dabei einerseits

- der im Projekt „Deutscher Wortschatz“ entwickelte Referenzcorpus aus derzeit ca. 300 Millionen laufenden Wortformen, sowie die in ihm enthaltenen zusätzlichen Informationen (cf. [Quasthoff & Wolff 00] [Heyer, Quasthoff, Wolff 02] und <http://wortschatz.uni-leipzig.de>), andererseits
- die im Rahmen dieses Vorhabens entwickelten Software-Werkzeuge zur Textanalyse, die grundsätzlich auf Textkollektionen beliebigen Umfangs angewandt werden können und als Analyseergebnis eine Datenbank aufbauen, über die (wenigstens) auf Basisdaten wie Wortfrequenzen und Frequenzklassen, Kollokationen oder Grundformrelationen zugegriffen werden kann (cf. [Heyer et al. 01a], [Heyer et al. 01b]).

<i>Fachbegriff</i>	<i>Häufigkeitsklasse Im Fachcorpus</i>	<i>Häufigkeitsklasse im Allgemeinsprachlichen Corpus</i>	<i>Differenz</i>
Hubraum	6	14	8
Nockenwelle	9	18	9
Fahrgeräusch	11	19	8
Zylinder	8	13	5

Tabelle 1: Frequenzvergleich von Fachbegriffen

Der bekannten These folgend, demzufolge sehr seltene Begriffe aufgrund ihrer zu hohen Spezifik für die Beschlagwortung ebenso wenig geeignet sind wie sehr häufige Begriffe (cf. [Salton 83: 62, insb. Abb. 3-2]), spielt bei diesem Ansatz zunächst der Vergleich von Frequenzklassen für Fachbegriffe eine wichtige Rolle bei der Auswahl von Kandidaten für die Textbeschlagwortung: Sowohl in Fach- als auch in Normcorpus hat jeder auftretende Begriff eine absolute sowie eine aus ihr im Verhältnis zur Corpusgröße berechnete relative Frequenzklasse.¹ Über die Voraussetzung einer Mindestfrequenzklasse lassen sich zu häufige bzw. zu seltene Begriffe aus der Analyse ausschließen, über den Vergleich von Frequenzklassen zwischen Fach- und Normcorpus lassen sich geeignete Beschlagwortungskandidaten finden, indem gefordert wird, dass als Kandidaten nur solche Wörter ausgewählt werden, deren Frequenzklasse im Fachcorpus wenigstens um eine Minstdifferenz verfügen (eine Differenz von 2 besagt dabei eine vierfache relative Häufigkeit im Fachcorpus). Am Beispiel einiger Fachbegriffe aus der Automobiltechnik sei dies verdeutlicht. Grundlage ist dabei ein Fachcorpus, das aus vier Jahrgängen einer bekannten Publikumszeitschrift zur Automobiltechnik besteht und das mit dem Normcorpus „Deutscher Wortschatz“ verglichen wurde (cf. [Wolff 01], [Heyer et al. 01b:81, insb. Tab. 7.7]):

¹ Die Frequenzklasse wird als logarithmisches Maß in Relation zum häufigsten Begriff eines Corpus ermittelt. Eine Klasse 4 besagt daher, dass ein Wort um den Faktor 16 (2^4) seltener gesehen wurde als das jeweils häufigste Wort im Corpus).

Der Vergleich von Häufigkeitsklassen ist ein vergleichsweise einfacher Ansatz, um die für einen Corpus charakteristischen Terme zu extrahieren. Kilgarriff sieht darüber hinausgehend in der Berechnung statistischer Prüftests für die in verschiedenen Corpora auftretenden Begriffe ein wesentliches Merkmal für die bessere Beschreibung von Corpora.²

2.2 Mehrschichtiger Corpusvergleich - ein Szenario für die Beschlagwortung von Texten

Für die praktische Anwendung des Corpusvergleichs auf das Problem der Beschlagwortung von Texten sei folgendes Szenario vorausgesetzt:

- Ein hinreichend großes Normcorpus steht als Referenzdatenbank zur Verfügung.
- Die zu beschlagwortende Textkollektion wächst im Vergleich zu ihrer Gesamtgröße relativ langsam.
- Jedes einzelne Dokument kann selbst als ein Textcorpus behandelt werden.

Die Grundlage des Beschlagwortungssystems ist zunächst eine Verallgemeinerung des wortfrequenzbasierten Corpusvergleichs, da hier das jeweils zu beschlagwortende Dokument als dritte Analyseebene hinzukommt: Die statistische Analyse, durchgeführt mit der im Umfeld des Projekts „Deutscher Wortschatz“ entwickelten Text Mining-Engine *Concept Composer* (cf. [Heyer, Quasthoff, Wolff 00], [Quasthoff & Wolff 00] u. unten Abb. 1), erfolgt für den Startbestand der Texte des jeweiligen Anwendungsgebietes sowie gesondert für jedes neue zu beschlagwortende Dokument. Zusammen mit der Corpusdatenbank des „Deutschen Wortschatzes“, die als *allgemeinsprachliche linguistische Datenbank* (ALDB) den Status eines Normcorpus aufweist, ergeben sich, anders als beim einfachen Corpusvergleich, für jedes Dokument zwei Vergleichsebenen:

- Vergleich mit den Analyseergebnissen für die aktuelle Dokumentenkollektion und
- Vergleich mit dem Normcorpus.

² „Corpus linguistics lacks a vocabulary for talking, quantitatively, about similarities and differences between corpora. [...]. One way of describing differences between corpora is by highlighting the words which have consistently been used more in the one corpus than the other“. [Kilgarriff 01: Kap. 10 – Conclusion].

Es ist offensichtlich, dass ein solcher frequenzbasierter Mehrebenenvergleich sprachliche Variation wie Vollformen, Schreibvarianten (Rechtschreibreform!), Gebrauch von Synonymen oder die Erkennung von Mehrwortbegriffen nicht berücksichtigen sollte. Deshalb tritt für die Auswahl von Beschlagwortungstermen eine zweite wesentliche Systemkomponente: Die *Expansion* und *Reduktion* der durch den mehrschichtigen Corpusvergleich ausgewählten Begriffe durch Rückgriff auf in der allgemeinsprachlichen linguistischen Datenbank vorhandenes linguistisches Wissen, z. B. über Vollform-/Grundformbeziehungen, Synonyme oder Eigennamen.

Als Ergebnis dieser Konzeption steht die Entwicklung eines modularen und hochparametrischen Beschlagwortungssystems, dessen Aufbau und Arbeitsweise im folgenden Kapitel näher beschrieben sind.

3 Systemarchitektur und Arbeitsweise

Das System ist als Web Service realisiert und kann über einen Webserver angesprochen werden. Dabei liegt ein einfaches Kommunikationsschema zugrunde: Der zu beschlagwortende Text wird per http an den Beschlagwortungsservice gesandt, dort wird die Textanalyse durchgeführt und das Beschlagwortungssystem liefert Schlagworte als nach Relevanz geordnete Liste an den Client, z. B. ein Content Management-System (CMS), zurück. Die Anzahl gewünschter Schlagworte kann dabei gesteuert werden. Abbildung 1 gibt hierzu einen Überblick. Das Kommunikations- sowie das nachfolgend beschriebene Steuerungsmodul stellen Querschnittkomponenten des Systems dar, die die einzelnen Module zu

- Textsegmentierung,
- Stoppworteliminierung und Frequenzabgleich,
- Expansion durch linguistisches Wissen und
- Ranking.

steuern.

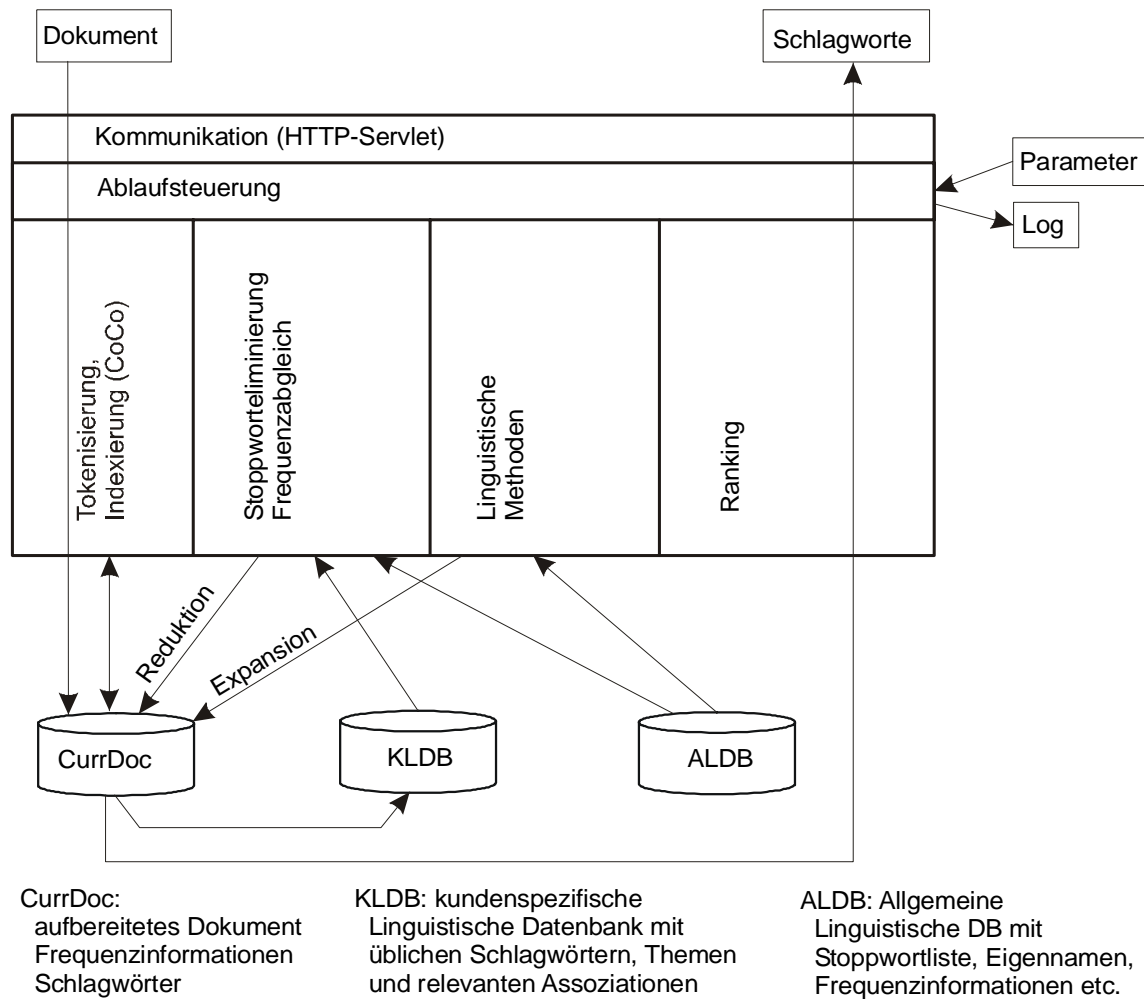


Abbildung 1: Systemarchitektur Concept Extractor

3.1 Steuerung

Das Steuerungsmodul stellt den einzelnen Komponenten des Beschlagwortungssystems einen Zugriff auf

- den aktuell zu verschlagwortenden Text,
- die in den verschiedenen Ebenen der Textanalyse ermittelten linguistischen und statistischen Daten.
- die globalen und komponentenspezifischen Parameter sowie
- ein Logging-System bereit.

Vom Steuerungssystem werden die einzelnen Module zur Bearbeitung der Daten angesteuert und ihre gegenseitigen Abhängigkeiten sichergestellt. Neben der Datenbank *CurrDoc* (Ergebnisse der Textanalyse für das aktuelle Dokument) wird von den Modulen auf eine allgemeinsprachliche Referenz-

datenbank (ALDB) mit umfangreichen linguistischen Daten und eine Datenbank mit dem gesamten Textbestand des Kunden (KLDB) zugegriffen. In dieser sind kundenspezifische Daten (Stoppwörter, Negativ- und Positivliste für Schlagwörter etc.) hinterlegt; sie wird fortlaufend um die neu verschlagworteten Texte ergänzt.

Die Module des Verschlagwortungsprozesses unterteilen sich nach ihren Aufgaben in die Bereiche

- Datenaufbereitung,
- Reduktion und Expansion,
- Bewertung.

Sie werden im Folgenden einzeln beschrieben.

3.2 Segmentierung, Erkennung von Mehrwortbegriffen, Häufigkeitsermittlung, Stoppwörter

Zur Tokenisierung des zu verschlagwortenden Textes wird die Text Mining-Engine *Concept Composer* verwendet. Sie zerlegt den Text in Sätze und Wörter und zählt Wörter aus. Sie enthält zudem eine automatische Erkennung von Mehrwortbegriffen auf der Basis einer umfangreichen, teilautomatisch generierten Mehrwortbegriffsliste des Deutschen, die durch Mehrwortbegriffe aus dem aktuellen Datenbestand der Anwendungsdomäne ergänzt wird (KLDB). In die Analyse gehen auch (HTML-)Strukturmerkmale ein, um eine höhere Gewichtung von Begriffen in Überschriften etc. gewährleisten zu können. Weitere relevante Layoutmerkmale wie etwa Großschreibung können berücksichtigt werden (jeweils durch Ersetzung der häufigsten Schreibvariante, z. B. ALTERNATIVMEDIZIN → Alternativmedizin).

Die aus dem „Deutschen Wortschatz“ gewonnene frequenzbasierte Stoppwortliste wird erweitert durch Stoppwörter aus der aktuellen Anwendungsdomäne. Sie wird zusätzlich um Großschreibungen und gebeugte Formen von Stoppwörtern ergänzt. Das entsprechende Modul entfernt in der *CurrDoc*-Datenbank alle Begriffe, für die ein Stoppwortflag in der Referenzdatenbank (KLDB) gesetzt ist.

3.3 Expansions- und Reduktionsmodule

Neben den im Text tatsächlich gebrauchten Wörtern kommen weitere Begriffe als Schlagwörter in Frage, die aus den verwendeten Begriffen hergeleitet

werden können. Im Einzelnen kommen dabei die nachfolgend beschriebenen Module zum Einsatz.

3.3.1 Schreibvarianten

Zu allen Wörtern aus dem Ausgangstext wird geprüft, ob es sich um Tippfehler handeln könnte. Dazu wird überprüft, ob das Wort in einer Liste korrekter Wörter enthalten ist. Ist das nicht der Fall, werden für längere Wörter typische Tippfehler mit Levenshtein-Abstand 1 generiert (z.B. Vertauschung zweier Buchstaben, Einfügen und Auslassen von Buchstaben, cf. [Navarro 01, Levenshtein 65]), und diese Varianten auf mögliche Korrektheit hin überprüft. Das Verfahren ist nicht für kurze Wörter anwendbar, da es für diese zu viele mögliche Varianten gibt, die gültige Wörter sind. Für die Wörter des Ausgangstextes wird zudem in der linguistischen Datenbank nachgeschlagen, ob sich die Schreibweise durch die Rechtschreibreform geändert hat. Wenn ja, wird die jeweils neue oder alte Variante ergänzt.

Synonymexpansion

Zu allen Wörtern des Ausgangstextes werden Synonyme aus der Synonymdatenbank der ALDB ergänzt. Diese sind jeweils niedriger bewertet als ihr Ausgangswort. Zusätzliche Informationen zur Stärke der Synonymierelationen können, falls vorhanden, zur differenzierten Gewichtung von Synonymen herangezogen werden.

Grundform- und Wortartbehandlung

Anhand der Informationen in der ALDB werden zu den Wörtern des Ausgangstextes Grundformen ermittelt. Dabei werden

- für gebeugte *Nomina* die Grundform *ergänzt* (Expansion),
- gebeugte *Adjektive* und *Verben* durch ihre Grundform ersetzt, wenn die Grundform nicht deutlich niedrigerfrequent als die gebeugte Form ist.

Weiterhin erhalten *Nomina* eine höhere Bewertung für ihre Eignung als Schlagwörter als *Adjektive* und *Verben*.

3.4 Frequenzabgleich

Im Anschluss an die zuvor genannten Expansions- und Reduktionsschritte werden alle Wörter einem Frequenzvergleich mit dem in der ALDB gespeicherten Referenzkorpus unterzogen³. Hierbei sind zwei Probleme zu berücksichtigen: das zur korrekten Berechnung der Häufigkeitsklasse notwendige

³ Ein Frequenzvergleich mit der KLDB ist im hier beschriebenen Prototypen noch nicht implementiert.

häufigste Wort des Referenzkorpus kann im zu verschlagwortenden Dokument *CurrDoc* fehlen, und nicht alle Worte aus *CurrDoc* kommen im Referenzkorpus vor. Solche Worte dürfen jedoch nicht a priori ignoriert werden, weil es sich dabei um als Schlagworte relevante Eigennamen handeln kann. Im Prototypen kommt ein Häufigkeits-Quotient zum Einsatz, dessen Logarithmus eine Näherungsformel für die Differenz der Häufigkeitsklassen darstellt, der aber die angesprochenen Probleme vermeidet. Der in der Reduktions- / Expansions-Phase berechnete Score wird dabei gewichtet mit dem Faktor:

$$\#(CurrDoc, Wort) / (\#(ALDB, Wort) + \#(CurrDoc, Wort))$$

wobei $\#(CurrDoc, Wort)$ bzw. $\#(ALDB, Wort)$ die absolute Häufigkeit von *Wort* im zu verschlagwortenden Text bzw. im Referenzkorpus *ALDB* ist. Dadurch werden Worte bevorzugt, welche in der *ALDB* selten vorkommen, insbesondere, wenn sie in *CurrDoc* mehrfach auftreten. Der Score von Worten, welche in der *ALDB* nicht vorkommen, bleibt nach dieser Formel unverändert. Worte, welche in der *ALDB* häufig vorkommen, werden dagegen stark abgewertet.

3.5 Heuristiken

Anschließend an den Frequenzabgleich werden verschiedene Heuristiken angewandt, um gute Schlagwortkandidaten mit einem besonderen Bonus zu versehen.

Eigennamen

Gute Kandidaten für Schlagwörter sind Eigennamen (*named entities*). Zu deren Erkennung kann auf einen umfangreichen Bestand geographischer Namen, Namen von Persönlichkeiten und Bezeichnungen von Pflanzen und Tieren in der *ALDB* zurückgegriffen werden. Außerdem erhalten Begriffe eine höhere Wertung, die als *Sachgebietsbezeichnung* gebräuchlich oder als *Personennamen* bekannt sind. Diese Informationen können um Informationen aus dem Textbestand der Anwendungsdomäne ergänzt werden.

HTML-Markup

Worte, die durch das HTML-Layout besonders hervorgehoben sind (z.B. Link-Texte, Überschriften, fette oder kursive Schrift etc.), werden als wichtig erkannt.

Manuelle Schlagwörter

Besonders gute und einfach zu ermittelnde Kandidaten für Schlagwörter stellen bereits früher verwendete, manuell vergebene, Schlagwörter dar. Eine Liste solcher Schlagwörter wird in der KLDB vorgehalten.

3.6 Bewertung

Zur Erzeugung der Schlagwortliste wird die gewichtete Variante des Bagging-Algorithmus von Breiman [Breiman 94] verwendet. An einer zufällig gewählten, konstanten Menge von Testbeispielen wurden die verschiedenen Expansionsschritte entwickelt. Diese bewerten die vorhandenen Wörter und fügen neue Wörter mit einer Bewertung für ihre Tauglichkeit als Schlagwörter hinzu. Der Bewertungsalgorithmus fasst die einzelnen Resultate gewichtet zusammen. In Abhängigkeit der Bewertung und der Länge der zu erzeugenden Schlagwortliste werden die Schlagwörter nach Relevanz sortiert zurückgegeben.

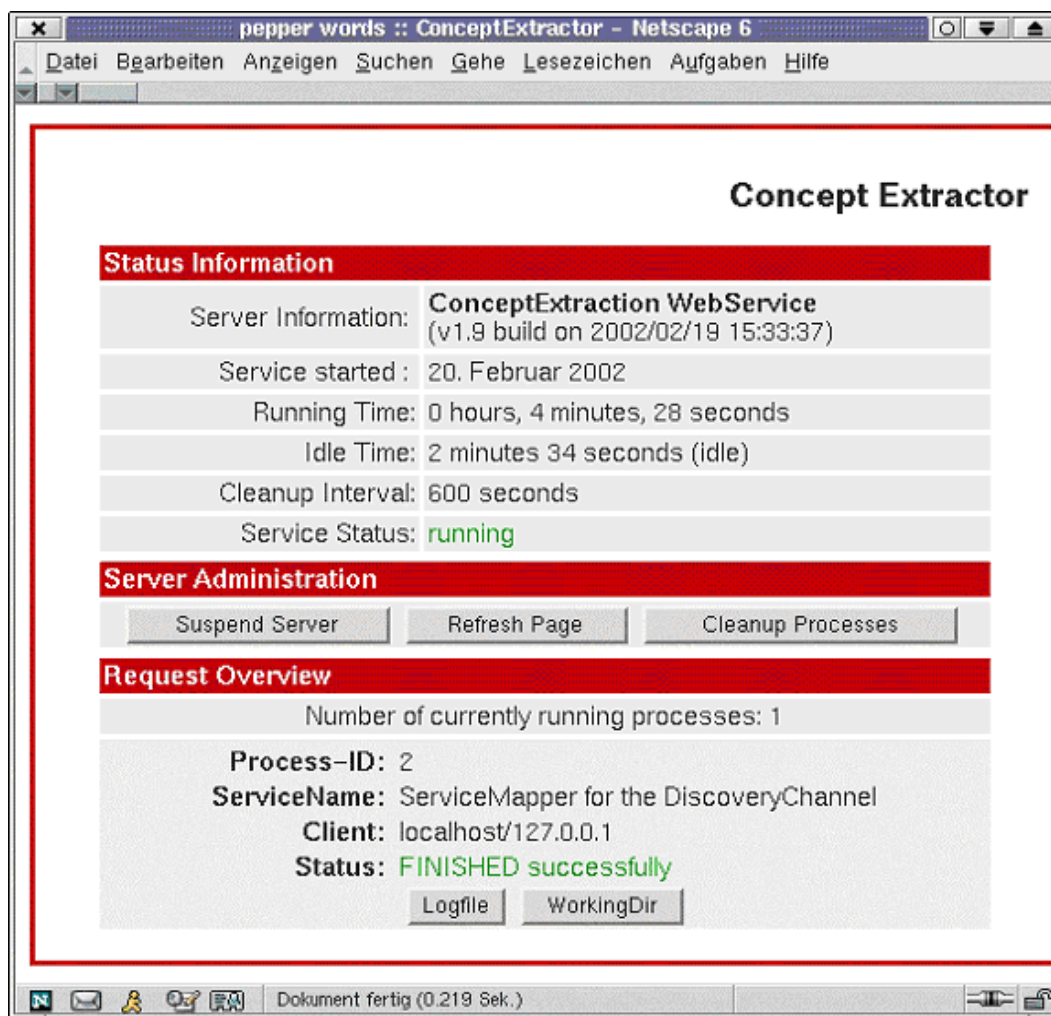


Abbildung 2: Webbasiertes Steuerinterface des Beschlagwortungsservers

3.7 Administration und Steuerung

Um das Beschlagwortungssystem als *web service* bereitstellen und steuern zu können, wurden zusätzlich serverseitige Komponenten entwickelt, die mit Hilfe von *Java Server Pages* (JSP) ein browserbasiertes Administrationsinterface generieren und so die Steuerung und Kontrolle des Beschlagwortungs-servers ermöglichen. Die Abbildungen 2 und 3 zeigen jeweils exemplarisch das Steuerinterface sowie die Logging-Ausgabe, die detaillierte Information zum Ablauf der Beschlagwortung enthält.

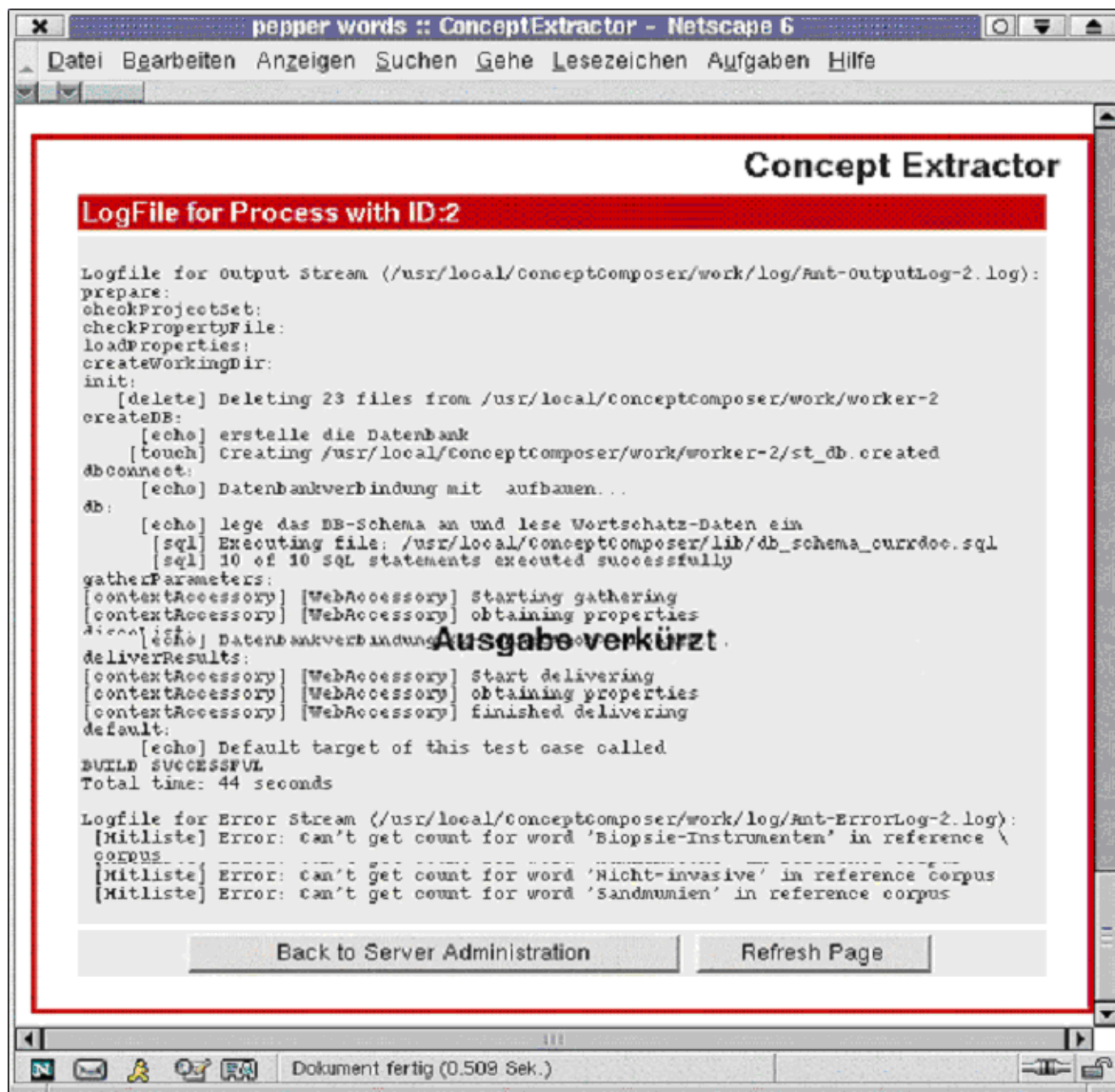


Abbildung 3: Webbasiertes Logging für den Analyseprozess des Concept Extractor

Zusätzlich zu obigen Steuerungsmechanismen existiert ein Testinterface, durch das mit einer webbasierten Upload-Funktion ein beliebiger Text über einen Webbrowser an den Beschlagwortungsserver geschickt werden kann.

4 Ein Beschlagwortungsbeispiel

Abschließend soll an einem Beispiel die Ergebnisqualität der automatischen Beschlagwortung verdeutlicht werden. Als Beispiel wird dabei der in Abb. 4 gezeigte Text über Alternativmedizin verwendet. Von der für den Text verantwortlichen Online-Redaktion wurden dabei durch manuelle Beschlagwortung folgende Begriffe vergeben:

Medizin, Alternative Heilmethoden, Homöopathie, Traditionelle Chinesische Medizin, Akupunktur, Shiatsu, Ayurveda

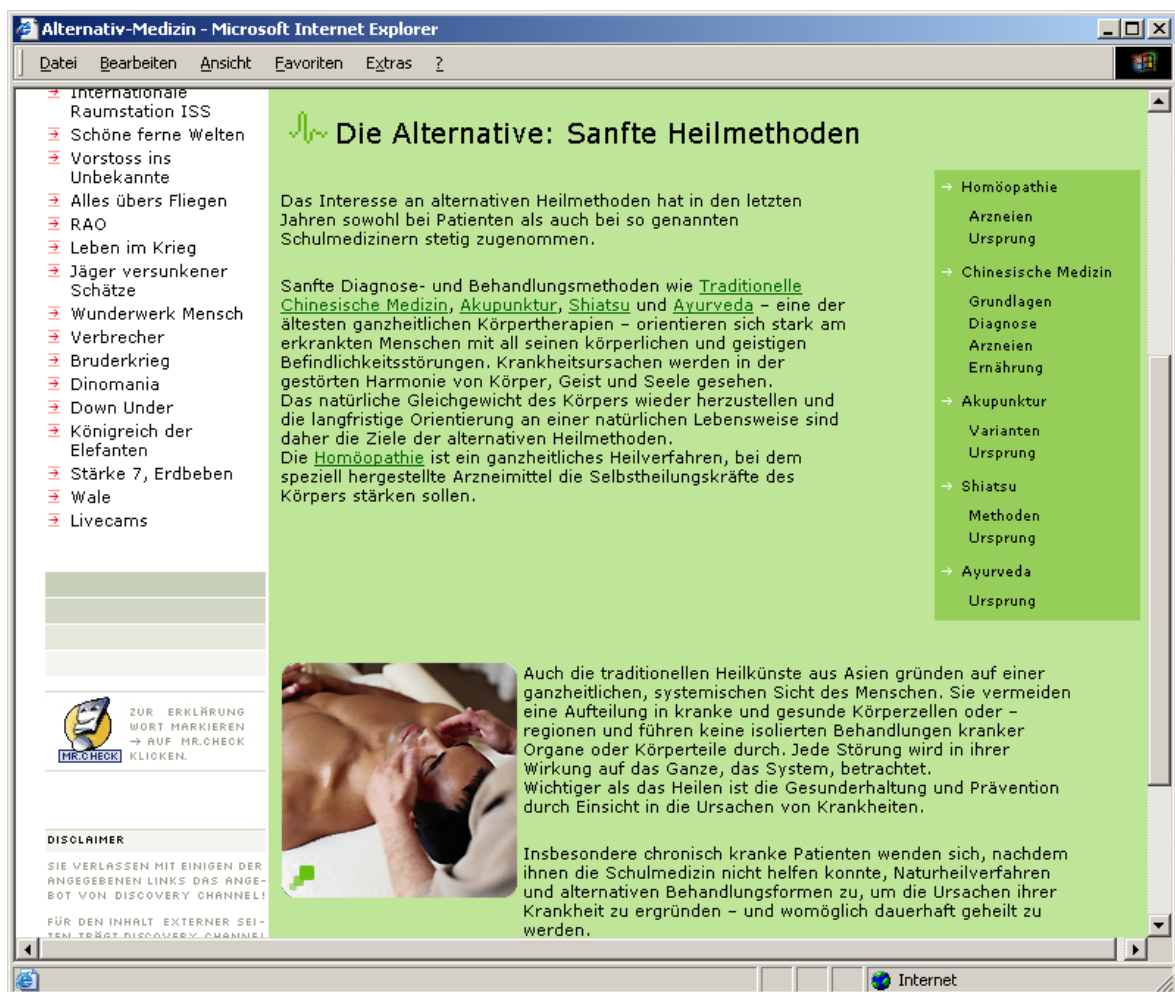


Abbildung 4: Beispieltext Alternativmedizin

Mit Hilfe des Concept Extrator wurden die nachfolgenden Schlagworte ermittelt - die Ausgabe ist dabei nach dem Wert der Zuordnungsfunktion gerankt; die fett gedruckten Begriffe stellen dabei die Untermenge dar, die als erste

256 Zeichen zur Aufnahme in die HTML-Meta Tags der Webseite vorgesehen sind, um die Erfassung durch Suchmaschinen zu verbessern:

Alternativ-Medizin, Shiatsu, Ayurveda, Medizin, Special, Akupunktur, Heilmethoden, Traditionelle, Arzneien, Homöopathie, Chinesische, Heilkunst, Diagnose, Ernährung, Körpertherapie, Grundlagen, Heilmethode, Heilkünste, Körpertherapien, Schulmediziner, Befindlichkeitsstörung, Sanft, Krankheitsursache, Behandlungsform, Störung, Organe, Ursprung, Gleichgewicht, Schulmedizinern, Selbstheilungskraft, Befindlichkeitsstörungen, Orientierung, Gesunderhaltung, Sanfte, Krankheitsursachen, Behandlungsformen, Krankheiten, Seele, Geist, Asien, Selbstheilungskräfte, Behandlungsmethode, Körperzelle, Heilverfahren, Krankheit

Das Beschlagwortungsergebnis zeigt, wie sich mit Hilfe frequenzbasierten Corpusvergleichs und unter Heranziehung zusätzlicher Wissensmodule eine flexible Beschlagwortung erreichen lässt. Dabei sind allerdings auch Schwächen des Systems offensichtlich: Nicht alle Mehrwortbegriffe können automatisch als solche erkannt werden (z. B. Aufnahme des Adjektivs chinesische als Schlagwort, Fehler bei der Groß- und Kleinschreibung).

5 Fazit

Das vorgestellte Beschlagwortungssystem stellt einen Mittelweg zwischen vollautomatischer Volltextindexierung und manueller Vergabe von Schlagworten dar. Durch Corpusvergleich kann die Schlagwortselektion an die Eigenheiten der jeweiligen Dokumentkollektion angepasst werden, zusätzliche Module für die Begriffsexpansion und –reduktion gewährleisten, dass auch sprachliche Varianten in die Beschlagwortung aufgenommen werden. Die Flexibilität des Systems lässt es für sehr unterschiedliche Probleme der Informationserschließung, gerade auch für innovative Informationsdienste aus dem Umfeld des *mobile computing* geeignet erscheinen, für die Restriktionen z. B. bezüglich der möglichen Beschlagwortungstiefe gelten.

Die systematische Evaluierung des Ansatzes steht noch aus. Bislang erfolgte nur eine qualitative Bewertung durch Online-Redakteure in einer ersten Anwendungsdomäne des Systems im Bereich Wissenschaftsjournalismus.

Literaturverzeichnis

- [Atkins et al. 02] Atkins, S. et al. (2002). „From Resources to Applications. Designing the Multilingual ISLE Lexical Entry.“. In: Proc. LREC-2002. Third International Conference on Language Resources and Evaluation. Las Palmas, May 2002, Vol. II, 687-692.
- [Baeza-Yates & Ribeiro-Neto 99] Baeza-Yates, R.; Ribeiro-Neto, B. (1999). Modern Information Retrieval. Harlow et al.: The ACM Press/The MIT Press.
- [Breiman 94] Breiman, L. (1994). „Bagging Predictors“. Technical Report No. 421, University of California at Berkeley, Department of Statistics, September 1994.
- [Heyer et al. 01a] Heyer, G.; Läuter, M.; Quasthoff, U.; Wittig, Th.; Wolff, Ch. (2001). „Learning Relations Using Collocations“. In: Proc. IJCAI Workshop on Ontology Learning, Seattle/WA, August 2001, 19-24.
- [Heyer et al. 01b] Heyer, G.; Läuter, M.; Quasthoff, U.; Wolff, Ch. (2001): „Wissensextraktion durch linguistisches Postprocessing bei der Corpusanalyse“. In: Lobin, H. (ed.) (2001). Sprach- und Texttechnologie in digitalen Medien. Proc. GLDV-Jahrestagung 2001, Universität Gießen, 71-83
- [Heyer, Quasthoff, Wolff 00] Heyer, G.; Quasthoff, U.; Wolff, Ch. (2000). „Aiding Web Searches by Statistical Classification Tools“. In: Proc. Proc. 7. Intern. Symposium f. Informationswissenschaft ISI 2000, UVK, Konstanz, 163-177.
- [Heyer, Quasthoff, Wolff 02] Heyer, G.; Quasthoff, U.; Wolff, Ch. (2002). „Knowledge Extraction from Text: Using Filters on Collocation Sets.“ In: Proc. LREC-2002. Third International Conference on Language Resources and Evaluation. Las Palmas, May 2002, Vol. III, 241-246.
- [Kilgarriff 01] Kilgarriff, Adam (2001). „Comparing Corpora“ In: International Journal of Corpus Linguistics 6(1) (2001), 97-133.
- [Levenshtein 65] Levenshtein, V. (1965). „Binary Codes Capable of Correcting Spurious Insertions and Deletions of Ones“. In: Probl. Inf. Transmission 1 (1965), 8–17.
- [Navarro 01] Navarro, G. (2001). „A Guided Tour to Approximative String Matching“. In: ACM Computing Surveys 33(1) (2001), 33-88.
- [Preece & Decker 02] Preece, A.; Decker, M. (2002). „Intelligent Web Services“. In: Intelligent Systems 17(1) (2002), 15-17.
- [Quasthoff & Wolff 00] Quasthoff, U.; Wolff, Ch. (2000). „An Infrastructure for Corpus-Based Monolingual Dictionaries“. In: Proc. LREC-2000. Second International Conference on Language Resources and Evaluation. Athens, May / June 2000, Vol. I, 241-246.
- [Rayson & Garside 00] Rayson, P.; Garside, R. (2000). „Comparing Corpora Using Frequency Profiling“. In: Proc. Workshop on Comparing Corpora, held in conjunction with the 38th annual meeting of the Association for Computational Linguistics (ACL 2000), Hong Kong, October 2000, 1-6.
- [Salton & McGill 83] Salton, G.; McGill, M. J. (1983). Introduction to Modern Information Retrieval. New York et al.: McGraw-Hill.
- [Simon 02] Simon, M. (2002). „Eine Botschaft auf allen Kanälen“. In: <e>Market, Juni 2002, 10-12.

[Wolff 01]. Wolff, Ch. (2001). „Aspekte des Vergleichs von Fach- und Normcorpora am Beispiel eines Fachcorpus aus der Automobiltechnik“. Arbeitsmaterialie, Universität Leipzig, Institut für Informatik, Abt. Automatische Sprachverarbeitung, Juni 2001.



Automatische Thesauruserstellung und Query Expansion in einer E-Commerce-Anwendung

Libo Chen, Ulrich Thiel, Marcello L'Abbate

Fraunhofer Institut
Integrierte Publikations- und Informationssysteme
(IPSI)
Dolivostraße 25
D-64293 Darmstadt
{chen,thiel,labbate}@ipsi.fhg.de

Zusammenfassung

Diese Arbeit beschreibt eine Methode für den automatischen Aufbau eines Thesaurus auf der Basis von bereits existierenden Kategorien von Dokumenten. Ein neues Clusteringverfahren, das „Layer-Seeds Verfahren“, wird in der Arbeit vorgestellt. Es bearbeitet die Terme in einer bestimmten Kategorie und organisiert sie zu einer baumähnlichen hierarchischen Struktur. Anhand dieser Struktur kann in einem interaktiven Retrievalsystem auf der Basis eines intelligenten Agenten die Benutzeranfrage expandiert werden, um bessere Suchergebnisse zu erzielen. Eine Pilotanwendung des Verfahrens wurde im Rahmen eines E-Commerce-Systems entwickelt.

Abstract

This work describes a method for the automatic construction of a thesaurus based on existing categories of documents. A clustering algorithm, “the layer seeds method”, is introduced, which facilitates the automatic generation of a thesaurus reflecting the specific vocabulary occurring in a given collection of documents. We assume that the collection is partitioned into document categories. The clustering works on terms extracted from the documents in a certain category and organizes them in a tree-like hierarchical structure. In an interactive retrieval system based on an intelligent agent, we used an automatically generated thesaurus to expand the user's queries, in order to obtain better search results. A pilot application of the method was integrated into an e-commerce website.



1 Einleitung

Online-Shopsysteme mit integrierten Suchfunktionen werden mittlerweile bei den meisten E-Commerce-Anwendungen in Kundengeschäft eingesetzt. Die Qualität der Informationen, die von solchen Online-shops angeboten werden, spielt für den erfolgreichen Geschäftsablauf eine entscheidende Rolle. Im betrachteten Anwendungsfall ist daher die Vermeidung von irrelevanten Angeboten ein wichtiges Ziel. Aber auch die Bereitstellung einer sinnvollen Auswahl von Produkten aus einer Kategorie muss gewährleistet sein. Beide Ziele sind i.a. erreichbar durch die Formulierung elaborierter Datenanfragen, die - zu einer aktuellen Kategorie von Produkten - eine Menge von Begriffen umfassen, welche die potentiell relevanten Objekte charakterisieren. Mit anderen Worten, um IR-Systeme zur Suche von Produktangeboten einsetzen zu können, benötigen wir ein Verfahren, das die üblicherweise von Benutzern ad-hoc eingegebenen Anfragen automatisch erweitert, und so die Folgen von zu vieldeutigen oder einseitig fokussierten Benutzeranfragen unter Verwendung des Kontextes der aktuellen Kategorie kompensiert. Query Expansion zählt zu einer der wichtigsten Methoden, dieses Problem zu lösen. Diese Methode versucht, die von Benutzern eingegebenen Anfragen „richtiger“ zu formulieren, damit die Benutzerwünsche möglichst unter Verwendung des Vokabulars der Informationsquelle repräsentiert werden. Dazu werden normalerweise Fachwörterbücher oder Thesauri verwendet. Bei Online-Shopping Systemen wurde bisher noch wenig relevante Arbeit in dem Bereich Query Expansion geleistet. Im Vergleich zu anderen Informationssystemen weist ein Online-Shopping System jedoch einige positive Eigenschaften auf, z. B. gut strukturierte Informationen über Produkte (Produktkategorien) oder praktische Tools zum Auffinden der Produkte (eingebaute Volltext-Suchmaschine). Die Ausnutzung dieser Eigenschaften kann die Aufgabe des Information Retrieval erleichtern.

Das im Folgenden beschriebene Verfahren zielt darauf ab, Anfragen der Benutzer automatisch zu expandieren, damit befriedigendere Ergebnisse mit einer der Situation angemessenen Anzahl von Produktangeboten ausgegeben werden. Dazu wird ein Thesaurus auf der Basis von bereits existierenden Produktkategorien durch ein spezielles Clusteringverfahren („Layer-Seeds Verfahren“) automatisch aufgebaut und zur automatischen Expansion der Benutzeranfragen verwendet. Somit lässt sich die Performance des gesamten E-Commerce Systems deutlich verbessern.

In diesem Paper stellen wir zunächst einige verwandte Ansätze vor. Dann präsentieren wir den Aufbau des Thesaurus anhand des „Layer-Seeds Verfahren“

und dessen Anwendung im E-Commerce Projekt COGITO. Die Anwendung wird anschließend experimentell evaluiert. Der Ausblick auf geplante Weiterentwicklungen beschließt diesen Beitrag.

2 Relevante Forschungsansätze

In den letzten Jahren wurden zahlreiche Forschungsprojekte im Bereich Query Expansion auf der Basis von Thesauri durchgeführt. Die automatische Query Expansion ist relativ einfach zu konzipieren bzw. zu implementieren. Dabei nehmen Benutzer nicht an dem Prozess teil. Der Expansionsvorgang läuft im Hintergrund. Die Ergebnisse werden dem Benutzer erst am Ende des Prozesses geliefert. Zur automatischen Query Expansion zählen z.B. die Arbeit von Spark-Jones [Spa71] und Crouch & Yong [CrY92]. Minker et al. [Mink72] zweifelten jedoch an dem Ergebnis von Spark-Jones. Ihre Forschung zeigte, dass in manchen Fällen die Termklassifikation und die darauf basierende Query Expansion zu einer Effizienzminderung des Information Retrieval führen könnten. Salton [Salton73], in einem Review über die Arbeit von Minker et al., hat ihren experimentellen Ansatz in Frage gestellt, somit blieb die Problematik der Effektivität der automatischen Query Expansion immer noch offen und muss bis heute im Einzelfall empirisch nachgewiesen werden. Eine Alternative bieten interaktive Verfahren wie das Relevance Feedback, bei dem Benutzer ihre Anfragen angeben und entscheiden, welche Dokumenten unter den Ergebnissen relevant sind. In diesen relevanten Dokumenten wird dann nach Termen gesucht, die später für Query Expansion verwendet werden [Rocchio71]. Eine andere Möglichkeit ist, dass den Benutzern eine Liste von Termkandidaten geliefert wird und sie treffen die Entscheidung, welche Terme in die Query aufgenommen werden sollen [Ekm92].

Grob gesehen ist ein Thesaurus eine Menge von Termen plus einer Menge von Relationen zwischen diesen Termen. Bisher werden die meisten Thesauri manuell, in einer zunehmenden Anzahl von Projekten aber bereits automatisch aufgebaut. Die manuell erstellten Thesauri lassen sich wiederum in zwei Typen unterteilen. Der erste umfasst die Universalthesauri wie Roget's und WordNet, die Bedeutungsrelationen wie Antonymie und Synonymie für ein einzelnes Thesauruselement enthalten. Sie werden selten in IR-Systemen verwendet. Der zweite betrifft IR-orientierte Thesauri wie INSPEC, die normalerweise Relationen unter Elementen im Thesaurus enthalten, wie z.B. BT (Broader Term), NT (Narrow Term), UF (Used For), and RT (Related To). Diese Art Thesaurus wird allgemein in kommerziellen Systemen verwendet. Ein Thesaurus kann auch automatisch berechnet werden, indem man in großen Dokumentsammlungen aus dem gemeinsamen Auftreten von Wörtern in

Dokumenten oder Dokumentteilen assoziative Beziehungen zwischen ihnen ableitet [Spa71; CrY92; Qiu93]. Die beiden Aufbaumethoden haben ihre Vor- und Nachteile. Die manuell erstellten Thesauri weisen zwar sehr gute Elemente und Struktur auf, sind aber sehr aufwendig zu konstruieren bzw. zu pflegen. Die automatisch generierten Thesauri dagegen haben zwar nicht immer die richtigen Elemente und meist eine Struktur, die deren semantische Beziehungen nur ungenau und fehlerhaft reflektiert, dafür sind sie aber einfacher und schneller zu erzeugen. Daher können sie aktueller sein und sowohl umfassend als auch für spezifische Fachgebiete bereitgestellt werden. Im folgenden Abschnitt wird versucht, eine Kombination der beiden Thesaurusaufbaumethoden zu finden, die die Vorteile der beiden Methoden aufweist.

Panyr [Panyr86] hat in seiner Arbeit den Steinadler-Ansatz vorgestellt, bei dem sowohl Terme als auch Dokumente automatisch klassifiziert werden. Der Vorgang läuft in drei Phasen ab. Zuerst werden Terme anhand ihrer absoluten Dokumentshäufigkeiten im ganzen Textcorpus in mehreren Prioritätsklassen eingeteilt. Anschließend wird in der einzelnen Klasse eine Clusteranalyse durchgeführt, damit Terme und Dokumente in Clusters organisiert werden. Dabei ist irrelevant, welches Clusterverfahren angewandt wird. Zum Schluss werden die benachbarten Prioritätsklassen anhand der gemeinsamen Dokumentenclusters miteinander verglichen und verknüpft.

Ein ganz anderer Ansatz geht in die Richtung automatische hierarchische Text-Kategorisierung [Göv99; Jicr94; From01]. Ein beliebiges Dokument wird unter Berücksichtigung der Struktur von existierenden Kategorien einer zu ihm am besten passenden Kategorie zugeordnet. Dabei wird eine Kategorie als Megadokument betrachtet und weiterhin als Termvektor repräsentiert, der mit dem Termvektor des zu klassifizierenden Dokuments verglichen wird. Die hierarchische Struktur von Kategorien wird dadurch berücksichtigt, dass die Beziehungen unter Kategorien gewichtet und in die Gewichtung der Ähnlichkeit zwischen dem Dokument und den Kategorien einbezogen werden. Ein Clustering von Termen oder Dokumenten findet bei dieser Methode nicht statt.

3 Automatischer Thesaurusaufbau auf der Basis des „Layer-Seeds Verfahren“

Der automatische Vorgang des Thesaurusaufbaus basiert auf manuell erstellten bzw. gepflegten Kategorien. Sie sind Repräsentation des menschlichen Wissens der Anwendungsexperten, die durch die Auswahl und Zusammen-

stellung eines der Anwendung angepassten Kategoriensystems den semantischen und z.T. auch pragmatischen Kontext vorgeben. So ist es nicht überraschend, wenn z.B. die konkreten Kategorien in einer Buchhandelsanwendung von den im Bibliothekswesen gebräuchlichen Klassifikationen etwas abweichen. Wir gehen davon aus, dass die Dokumentkollektion anhand dieser Kategorien sortiert ist, wodurch bereits eine Zuordnung zwischen einer Kategorie und den Termen der in ihr enthaltenen Texte induziert wird. Der im Folgenden beschriebene Vorgang wird nun jeweils nur auf Terme aus Texten einer spezifischen Kategorie d.h. auf eine viel kleinere Datenmenge als die Gesamtmenge angewendet. Dies hat zur Folge, dass die Präzision der Ergebnisse stark erhöht und der Rechenaufwand enorm reduziert wird. Manche komplizierte Clustering Methoden werden erst dadurch realisierbar.

Der eigentliche Vorgang des Aufbaus lässt sich hauptsächlich in 3 aufeinanderfolgenden Phasen untergliedern: Kollektion bzw. Auswahl der Thesauruselemente, Beziehungsfeststellung zwischen den Thesauruselementen und Klassifizierung der Thesauruselemente anhand der Beziehungen. Zu der letzten Phase wird in dieser Arbeit eine Methode mit dem Namen „Layer-Seeds Verfahren“ konzipiert.

Um die Elemente des Thesaurus bestimmen zu können, werden zuerst Terme aus allen Daten unter einer bestimmten Kategorie gesammelt. Sie werden anschließend normalisiert. Die Normalisierung der Terme umfasst zwei Schritte: Stoppworteliminierung und Grundformreduktion. Dabei sind die Vorgehensweisen je nach Sprache unterschiedlich. In dieser Arbeit sind die Normalisierungsmethoden auf die deutsche Sprache ausgerichtet. Nach der Normalisierung werden noch bestimmt, welche übriggebliebenen Terme in den Thesaurus einbezogen werden müssen. Die Entscheidung geht davon aus, dass, wenn ein Term in eine Kategorie miteinbezogen werden soll, er auch kategoriespezifisch sein muss. Die kategorie-übergreifenden Terme sollen somit gelöscht werden. Es wird ein Kriterium eingesetzt, um zu überprüfen, ob ein Term kategoriespezifisch ist oder nicht. Das Kriterium lautet: Wenn sich ein Term gleichverteilt in verschiedenen Kategorien befindet, ist dieser Term *nicht* kategoriespezifisch. Um die Verteilungen eines Terms in zwei Kategorien zu vergleichen, werden folgende Formeln verwendet, die in Analogie zum Vorgehen bei der Varianzanalyse entwickelt wurden. Gegeben sei ein Term, der schon normalisiert ist. Es wird angenommen, dass der Term in m Kategorien vorkommt mit den Häufigkeiten hk_1, hk_2, \dots, hk_m .

$$\text{Summe der Häufigkeiten: } S = \sum_{i=1}^m h k_i \quad \text{Mittelwert: } MW = \frac{S}{m}$$

$$\text{Verteilungsgrad} = \sum_{i=1}^m \left(\frac{h k_i - MW}{S} \right)^2 = \frac{1}{S^2} \sum_{i=1}^m (h k_i - MW)^2$$

Der Verteilungsgrad misst die Unterschiede zwischen den Häufigkeiten, mit denen der Term in jeweils verschiedenen Kategorien auftritt. Je kleiner der Wert, umso mehr ist der Term gleichverteilt. Wenn der Wert größer als ein bestimmter Schwellenwert ist (Schw1), so ist der Term nicht in allen Kategorien gleich verteilt und kann in den Thesaurus miteinbezogen werden. Dieser Verteilungsgrad unterscheidet sich von der Varianz dadurch, dass hier das Quadrat der Häufigkeitssumme S statt der Anzahl der Kategorien m als Nenner genommen wird. Somit werden die Häufigkeiten der Terme normalisiert, so dass ein einheitlicher Schwellenwert gefunden werden kann, um mit dem Verteilungsgrad die Ähnlichkeit der Verteilungen für alle Terme zu bestimmen.

Nachdem die Elemente für den Thesaurus bestimmt sind, werden die Beziehungen zwischen den Elementen festgestellt. Bei dem in dieser Arbeit zu erstellenden Thesaurus sind zwei Termbeziehungen zu betrachten, nämlich die „Thematische Spezialisierung“ und die „Relevant“-Beziehung. Alle anderen Terme, die keine dieser zwei Beziehungen besitzen, sind in Bezug auf einander irrelevant. Um festzustellen, welche Beziehung zwischen zwei konkreten Thesauruselementen besteht, werden zuerst die Terme mit Termvektoren repräsentiert. Das i -te Element im jeweiligen Termvektor ist die Häufigkeit des Terms im i -ten Dokument. Die Dimension des Vektors ist die Anzahl der Dokumente in einer Kategorie.

Das Kosinusmaß wird eingesetzt, um die Relevanz zwischen zwei Termen zu berechnen, wobei V_x, V_y zwei Termvektoren und X_k und Y_k das k -te Element im jeweiligen Termvektor sind. Das Kosinusmaß hat den Vorteil, nur den Winkel zwischen beiden Vektoren zu berücksichtigen. Dadurch wird der Einfluss des Längenunterschieds von Dokumenten eliminiert und die absolute Häufigkeit der Terme ist nicht mehr relevant.

$$\cos(V_x, V_y) = \frac{V_x \cdot V_y}{|V_x| \cdot |V_y|} = \frac{\sum_{k=1}^n X_k Y_k}{\sqrt{\sum_{k=1}^n X_k^2} \cdot \sqrt{\sum_{k=1}^n Y_k^2}}$$

Es wird noch ein Schwellenwert (Schw4) eingeführt. Zwei Terme, deren Kosinusmaß größer als dieser Schwellenwert ist, werden dann als wechselseitig

relevant betrachtet. Bei dieser Vorgehensweise wird jedoch eine mögliche Beziehung ignoriert. Wenn ein Term immer mit einem anderen Term zusammen in demselben Dokument in einer bestimmten Kategorie vorkommt, der zweite Term aber nur in einem viel geringeren Maß mit dem ersten Term zusammen auftritt, so ließe sich intuitiv behaupten, dass der zweite Term allgemeinere Bedeutung hat als der erste. Das heißt, es besteht eine „Thematische Spezialisierung“-Beziehung zwischen den beiden Termen. Diese Beziehung ist als unabhängig von der mit dem Kosinusmaß bestimmten „Relevant“-Beziehung anzusehen.

Um das Prinzip näher zu erläutern, werden hier Similaritätsgrade eingeführt. t_1 und t_2 seien Terme, $V(t_1)$ ist der Termvektor von Term1, $V(t_2)$ ist der Termvektor von Term2

Similarität 0: $s_0 = \cos (V(t_1), V(t_2))$

Similarität 1: $s_1 = \frac{\text{Anzahl der Dokumente, die } t_1 \text{ und } t_2 \text{ enthalten}}{\text{Anzahl der Dokumente, die } t_1 \text{ enthalten}}$

Similarität 2: $s_2 = \frac{\text{Anzahl der Dokumente, die } t_1 \text{ und } t_2 \text{ enthalten}}{\text{Anzahl der Dokumente, die } t_2 \text{ enthalten}}$

Mit einer großen Similarität s_0 ist die „Relevant“-Beziehung festzustellen. Eine kleine Similarität s_1 mit einer viel größeren s_2 bedeutet, dass t_2 relativ häufig mit t_1 zusammen vorkommt und t_1 nur wenig mit t_2 zusammen vorkommt. t_2 ist daher wahrscheinlich der Spezialisierungsbegriff von t_1 . Wie klein s_1 bzw. wie groß s_2 sein soll, um die „Thematische Spezialisierung“ behaupten zu können, muss vorher genau angegeben werden. Zwei zusätzliche Schwellenwerte ($Schw_2$ und $Schw_3$) werden daher eingeführt.

Mit dieser drei Similaritäten lassen sich Beziehungen zwischen Termen feststellen.

„Thematische Spezialisierung“:

$$\text{Wenn } \begin{cases} s_1 \leq Schw_2 \\ s_2 \geq Schw_3 \end{cases}$$

dann ist t_2 ein Spezialisierungsbegriff von t_1 . Diese Beziehung zeigt an, dass ein Thema – gegeben durch Begriff t_1 - durch den Begriff t_2 spezialisiert oder

eingeschränkt werden kann. Dies ist z.B. der Fall, wenn t2 ein Unterbegriff, eine Eigenschaft, ein Teil, oder auch nur ein stark assoziierter Begriff in Bezug auf t1 ist.

„Relevant“-Beziehung:

Wenn $\begin{cases} s1 > Schw2 \\ s0 \geq Schw4 \end{cases}$ O-
der $\begin{cases} s2 < Schw3 \\ s0 \geq Schw4 \end{cases}$
dann ist t1 für t2 relevant.

„Irrelevant“:

Wenn $\begin{cases} s1 > Schw2 \\ s0 < Schw4 \end{cases}$ oder $\begin{cases} s2 < Schw3 \\ s0 < Schw4 \end{cases}$
dann sind t1 und t2 irrelevant.

Die Feststellung der Schwellenwerte muss durch Experimente empirisch vorgenommen werden. Dabei sollen verschiedene Schwellenwerte ausprobiert werden, um zu überprüfen, bei welchen die beste Thesaurusstruktur geliefert werden kann. Wir nehmen $P_{\text{element}}(\text{Schw1})$ als die Gütewahrscheinlichkeit für die Elementauswahl unter der Einsetzung von Schwellenwert1 und $P_{\text{beziehung}}(\text{Schw2}, \text{Schw3}, \text{Schw4})$ als die Gütewahrscheinlichkeit, die darstellt, wie gut die Struktur des Thesaurus unter der Einsetzung von den Schwellenwerten2-4 ist.

$$P_{\text{element}}(\text{Schw1}) = \frac{\text{Anzahl der beobachteten Terme, die richtig im Thesaurus sind}}{\text{Anzahl der gesamten beobachteten Terme}}$$

$$P_{\text{beziehung}}(\text{Schw2}, \text{Schw3}, \text{Schw4}) = \frac{\text{Anzahl der beobachteten Beziehungen, die richtig im Thesaurus sind}}{\text{Anzahl der gesamten beobachteten Beziehungen}}$$

Die Richtigkeit der Terme und der Beziehungen wird von Menschen rein empirisch beurteilt. Die Schwellenwerte 1-4 müssen dann diejenige Werte einnehmen, die die höchste Gütewahrscheinlichkeit liefern, wobei sie mindestens 50% sein sollte.

Anhand der Beziehungen werden Terme zum Schluss mit dem „Layer-Seeds Verfahren“ klassifiziert. Dabei wird zuerst jedem Term eine Gewichtung zugewiesen. Die Gewichtung wird nach folgender Formel berechnet:

K-TF-DF Gewichtung: $G(t, K_i) = THK(t, K_i) * DHK(t, K_i)$

Wobei t ein Term, $THK(t, K_i)$ die Häufigkeit von t in der Kategorie K_i und $DHK(t, K_i)$ die Anzahl der Dokumente in der Kategorie K_i , die den Term t enthalten. Diese Formel besagt, je häufiger ein Term in der Kategorie vorkommt, desto wichtiger ist er; je breiter der Term in der Kategorie verteilt ist, desto wichtiger ist er. Nach der Termnormalisierung und der Löschung der Kategorie-unspezifischen Terme sind nur diejenigen Terme geblieben, die wichtig bzw. Kategorie-spezifisch sind. Je größer die Gewichtung eines solchen Terms ist, desto besser kann er die Kategorie vertreten. Die Wichtigkeit (Gewichtung) eines Terms liegt also nicht mehr darin, wie gut er die Dokumente unterscheidet, sondern darin, wie gut er eine Kategorie vertritt, wobei die DHK eine wichtigere Rolle spielt als THK. Anhand der Gewichtungsformel wird jedes Mal ein Term als „Samen“ in der Kategorie ausgewählt, dessen Gewichtung am größten unter den Termen ist, die noch nicht Samen sind. Diese Samen werden dann als Ausgangspunkte der späteren Prozedur dienen.

Der Algorithmus vom „Layer-Seeds Verfahren“ wird im Folgenden aufgeführt:

- a. Am Anfang existiert nur eine Schicht, die alle Terme enthält.
- b. Ein Term in der aktuellen Schicht wird als Samen gewählt, der noch kein Samen und dessen K-TF-DF Gewichtung am größten in dieser Schicht ist.
- c. Die drei Similaritäten zwischen dem Samen und allen anderen Termen in derselben Schicht, die keine Samen sind, werden berechnet und deren Beziehungen werden somit festgestellt. Diejenigen Terme, die eine „Thematische Spezialisierung“-Beziehung mit dem Samen haben, werden als Spezialisierungsbegriffe in der nächsten Schicht diesem Samen zugeordnet. Diese Beziehungen werden als „Direkte Thematische Spezialisierung“, bezeichnet. Es wird noch überprüft, ob die Spezialisierungsbegriffe auch „Thematische Spezialisierung“-Beziehungen mit den Vorfahren dieses Samens haben; wenn ja, werden die Beziehungen mit den Vorfahren als „Indirekte Thematische Spezialisierung“ bezeichnet.

Diejenigen Terme, die eine „Relevant“-Beziehung mit dem Samen haben, werden nicht bearbeitet, die Beziehung wird jedoch protokolliert. Diejenigen Terme, die „irrelevant“ bzgl. des Samens sind, bleiben in derselben Schicht unbearbeitet.

- d. Falls noch Terme, die noch keine Samen sind, in der aktuellen Schicht gefunden werden können, wird Schritt b wiederholt. Im anderen Fall und wenn es noch untere Schichten gibt, wird die Schichtnummer um eins erhöht und der Schritt b wiederholt; ansonsten wird der Prozess beendet.

Abbildung 1 stellt den Vorgang des Verfahrens dar:

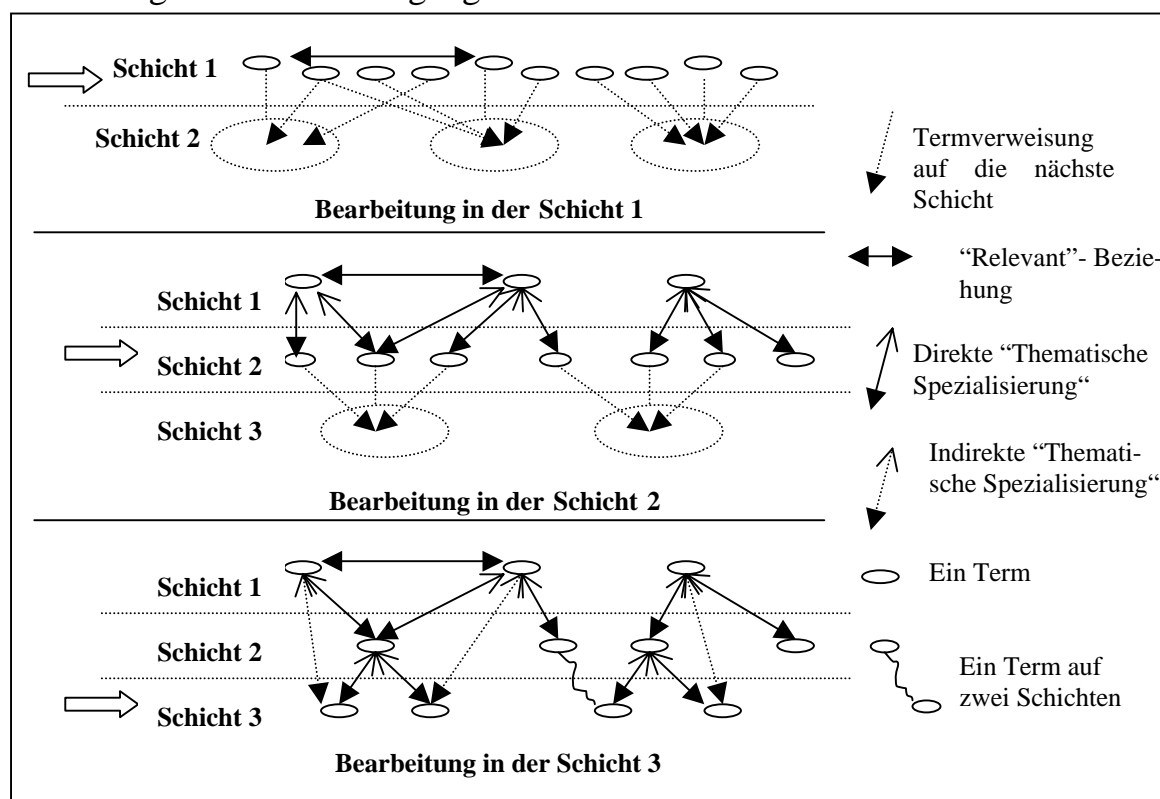


Abbildung 1: Vorgang des Thesaurusaufbaus nach dem „Layer-Seeds Verfahren“

Im Rahmen des COGITO Projektes (siehe Abschnitt 4) wurde das „Layer-Seeds Verfahren“ mit verschiedenen Schwellenwerten ausprobiert. Insgesamt 600 Terme bzw. Termpaare wurden ausgewählt und geprüft. Dabei wurden Kurzbeschreibungen von Computerfachbüchern als Dokumentmenge betrachtet. Tabelle 1 zeigt die Schwellenwerte, die die besten Ergebnisse liefern.

Schwellenwert	Wert	P
Schw1	0,1	0,8
Schw2	0,4	0,69
Schw3	0,8	
Schw4	0,05	

Tabelle 1: Schwellenwertfeststellung

Als Beispiel werden nun von zwei Kategorien jeweils 20 Samen in der ersten Schicht aufgeführt. Die Terme innerhalb einer Kategorie werden nach ihrer Gewichtung geordnet. Es ist ersichtlich, dass diese Terme angemessene Vertreter der jeweiligen Kategorie sind.

Kategorie Nummer: K1 Kategorie Name: Betriebssysteme

Samen: Windows, Linux, Microsoft, Arbeit, Unix, Lernen, Betriebssystem, Netzwerk, Konfiguration, Grundlage, Shell, Programmierung, Mac, Handbuch, Netzwerke, Multimedia, Umgang, einrichten, X, Software

Kategorie Nummer: K3 Kategorie Name: Informatik

Samen: Informatik, Grundlage, Software, Lehrbuch, Algorithmen, Methode, Computer, Sprache, Universität, C, Java, Wirtschaftsinformatik, methods, Multimedia, Übungsaufgaben, Based, Verständnis, Netz, bilden, Data

__1:Java

__2:Netzwerk

__2:Netscape

___3:Browser

___3:Ergebnis

___3:anstellen

___3:leben

___3:VBScript

___3:Stylesheets

___3:Dummy

___3:JavaScripts

___3:Informatio

___3:benutzer-
freundlichen

__2:Mail

__2:Enterprise

__2:Datenstrukturen

__2:Commerce

__2:Ausdrücke

__2:APIs

__2:Grundkonzepte

__2:Corba

__2:classes

__2:Aussagenlogik

__2:Webprogramming

__2:Unicode

__2:Threading

__2:RMI

___3:programming

___3:Security

___3:Objektorientierung

___3:JavaBeans

___3:Zugreifbarkeit

___3:Verteilung

___3:unternehmensweiter

___3:Typumwandlungen

___3:Toolkit

___3:Timecard

___3:Threadprogrammierung

___3:Technology

___3:Technologies

___3:Superklassen

___3:Subklassen

___3:Softwarelösungen

___3:Selection

___3:Resources

___3:Package

___3:OMG

___3:Methodenaufrufe

___3:Klassendeklarationen

___3:IDL

___3:Geltungsbereiche

__2:JDOM

__2:JDK

___3:Security

___3:Klassenbibliotheken

___3:Verteilung

___3:Java2D

___3:IDL

___3:Fibel

___3:Beispielanwendungen

__2:Applets

___3:Threads

___3:Security

___3:frames

___3:EDV

___3:Zugreifbarkeit

___3:Typumwandlungen

___3:Toolkit

___3:Threadprogrammierung

___3:Superklassen

___3:Suns

___3:Subklassen

___3:Methodenaufrufe

___3:Konvention

___3:Klassendeklarationen

___3:JFC

Abbildung 2: Spezialisierungsbegriffe von dem Term „Java“ in der Kategorie „Programmiersprachen“

Bei einem gegebenen Term können in dem Thesaurus seine allgemeineren Begriffe, Spezialisierungsbegriffe und relevante Begriffe nach angegebenen Relevanzstufen gefunden werden. Als Beispiele werden nachfolgend die Spezialisierungsbegriffe von dem Term „Java“ aus der Kategorie „Programmiersprachen“ genommen. Der Term „Java“ befindet sich in der ersten Schicht im Thesaurus. Es handelt sich um einen ziemlich breiten Begriff und er hat daher eine lange Liste von Spezialisierungsbegriffen. In der Abbildung 2 werden kleine Teile davon aufgelistet.

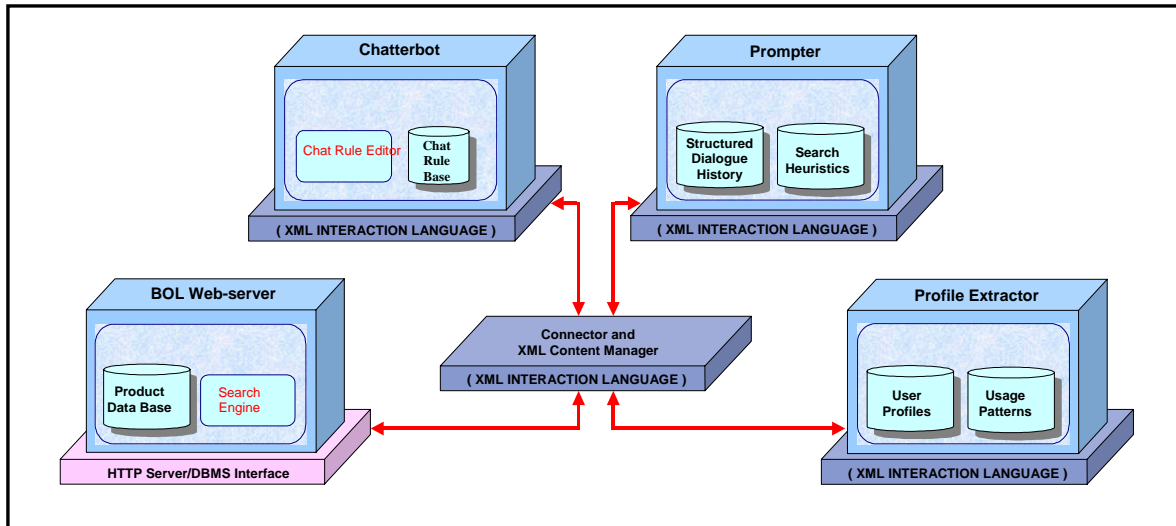
Es ist noch nicht klar, ob es universelle Schwellenwerte für alle Domänen gibt. Die für eine Domäne gültigen Schwellenwerte sind wahrscheinlich nicht auf andere Domäne übertragbar. Experimente müssen dann für jede einzelne Domäne durchgeführt werden.

4 Anwendung des Thesaurus

Ein konkretes Anwendungsbeispiel des Thesaurus und die darauf basierende automatische Query Expansion findet sich im EU-Projekt COGITO (Agenten für den E-Commerce mit personalisierter Interaktion) [Cogito], dessen Ziel die Bewältigung des Kommunikationsproblems im Bereich E-Commerce ist. In Zusammenarbeit mit BOL¹ wurden verschiedene Ansätze untersucht, um die Natürlichkeit, Einfachheit, Personalisierung und Effektivität der Kommunikation zu gewährleisten. Dabei wurde ein intelligenter Agent eingesetzt, der sich mit den Kunden nach vorher bestimmten Regeln in natürlicher Sprache unterhält (Chatterbot). Die Möglichkeiten heutiger Chatterbots werden im Projekt COGITO durch eine innovative Kombination moderner Techniken erweitert.

Dazu wird ein graphisches konversationales Interface implementiert, das nicht bloß auf Benutzereingaben reagiert, sondern in proaktiver Weise eine zielgerichtete Konversation mit dem Benutzer führen kann, um eine evtl. komplexe Aufgabe zu lösen. Unsicherheiten im Verständnis von Zielen/Aufgaben werden ausgeräumt, um den besten Lösungsweg zu finden, wobei die Reaktionen der Benutzer während des Dialogs erfasst und interpretiert werden. Auf dieser Basis kann das Verhalten des Systems adaptiert werden, z.B. durch Übernehmen der Initiative im Dialog, um neue Informationen oder Suchstrategien zu empfehlen. Die Natürlichkeit der Interaktion wird insbesondere für gelegent-

¹ Bertelsmann online. Internationaler Media- und Unterhaltungsshop im Internet <http://www.bol.com>.



liche Benutzer durch 2D- und 3D-Animationen erhöht, die im Bedarfsfall emotionale Reaktionen des Agenten simulieren.

Selbst „Chatterbots“ mit einem ausgefeilten Repertoire dialogischer Reaktionsmöglichkeiten bleiben jedoch nur auf ihren Unterhaltungswert beschränkt, wenn sie die Benutzer nicht als Individuen mit bestimmten Bedürfnissen, Vorlieben etc. wahrnehmen. In COGITO werden Verfahren zur inhaltsbasierten Filterung, bei denen Benutzerprofile inhaltliche Aspekte relevanter Dokumente oder Angebote erfassen, mit einem Ansatz zur kollaborativen Filterung kombiniert. Dieser Ansatz gruppiert Benutzer in Abhängigkeit ihrer Interessen und Motive und leitet daraus Empfehlungen innerhalb solcher virtuellen Gemeinschaften ab. Weil kollaborative Filterverfahren nur auf den von Benutzern geäußerten Meinungen basieren, sind sie besonders gut in Anwendungsfeldern einsetzbar, in denen der persönliche Geschmack eine große Rolle spielt, wie etwa bei der Auswahl von Belletristik, Spielfilmen oder Fernsehsendungen. Automatische Lernverfahren werden eingesetzt, um Benutzercharakterisierungen mit Vorlieben und Kaufentscheidungen zu assoziieren.

Abbildung 3: Die COGITO Systemarchitektur

Im Allgemeinen wird ein Benutzer nicht auf Anhieb geeignete Einträge in der Produktdatenbank finden. Um langwierige Abfragemodifikationen zu vermeiden, bezieht der Retrievalprozess in COGITO vier Komponenten ein (siehe Abbildung 3): den Chatterbot, das Content-Management der E-Commerce-Site, die Profilgenerierungsmaschine und ein auf den Benutzerprofilen basierendes Erweiterungssystem für Anfragen, „Prompter“ genannt, das dem Agenten Erweiterungsvorschläge für Datenanfragen in der Produktdatenbank „souffliert“. Der Prompter setzt einen Regelinterpreter ein, um die (XML-) Struktur der gesuchten Objekte oder Dokumente ebenso in die Präzisierung

des Suchprozesses einzubeziehen wie die Präferenzen der Benutzergruppen, die dem Profil des aktuellen Benutzers entsprechen.

Der Prompter realisiert außerdem den Zugriff auf den Thesaurus, der auf den BOL Bücherkategorien anhand des Layer-Seeds Verfahren aufgebaut ist, wobei Titel und Beschreibungen der Bücher als Dokumente für den Thesaurusaufbau verwendet werden. Basierend auf den gegenwärtigen Dialogbeiträgen und auf dem sich daraus ergebenden Kontext, wird vom Chatterbot automatisch eine aus Schlüsselwörtern bestehende Datenanfrage generiert. Diese wird vom Prompter expandiert, indem relevante Begriffe aus dem Thesaurus gewonnen werden. Im Folgenden wird dieser Prozess detaillierter aufgeführt.

Es wird angenommen, dass eine Anfrage aus n Stichwörtern besteht; für diese Anfrage ergeben sich m Ergebnisse; k ist die maximale Anzahl von Ergebnissen, die dargestellt werden sollen; k ist kleiner als m .

- a. Jedem Stichwort wird eine Kategorie zugewiesen, in der das Stichwort die größte Gewichtung hat.
- b. In der jeweiligen Kategorie werden für jedes Stichwort seine „relevanten“ Begriffe im Thesaurus gesucht. Falls mehrere Stichwörter gemeinsame „relevante“ Begriffe haben, werden die Similaritäten s_0 von ihnen zusammenaddiert. Die gefundenen „relevanten“ Begriffe werden in der Relevantbegriffliste gespeichert.
- c. Der Begriff, der zum Schluss den größten Wert s_0 besitzt, wird vom Thesaurus als Expansionswort vorgeschlagen.
- d. Dieses Wort wird mit der ursprünglichen Anfrage mit „AND“ verknüpft. Wenn die Anzahl der Ergebnisse gleich 0 ist, eignet sich das Wort nicht für die Anfrageexpansion. Es wird von der Relevantbegriffliste gelöscht. Schritt c wird wiederholt. Falls die Anzahl der Ergebnisse größer als 0 und kleiner gleich k ist, wird der Prozess beendet.

Die automatische Anfrageexpansion muss sehr vorsichtig angegangen werden, weil nur die Benutzer selbst genau wissen, welche Produkteigenschaften sie eigentlich suchen. Ohne Interaktion mit dem Benutzer ist es schwierig, unter verschiedenen Expansionsmöglichkeiten zu entscheiden. Die „Thematische Spezialisierung“-Beziehung im Thesaurus ist daher nicht zur automatischen Expansion verwendbar. Nur die Begriffe, wie z.B. der Name der Kategorie oder die relevanten Begriffe, können zur automatischen Expansion eingesetzt werden. Trotzdem konnten interessante Ergebnisse aus der prototypischen Anwendung des COGITO Systems erkannt werden. Diese sind im nachfolgenden Kapitel aufgeführt.

5 Evaluation

Die Evaluation eines Systems umfasst normalerweise drei prozedurale Ebenen:

- die Verifikation, bestehend aus der Überprüfung der Operations-Implementierung, die sich direkt aus den Benutzeranforderungen, sowie aus den operationalen Anforderungen ergibt;
- die Evaluierung, mit dem Ziel der Überprüfung der aus den Benutzeranforderungen und aus den prozeduralen Anforderungen abgeleiteten Funktionalitäten;
- und schließlich die Valutierung, welche eine Kontrolle des Systemnutzens anstrebt. Dabei wird untersucht, ob durch dessen Verwendung eine Steigerung der Effizienz und der Erfolgsrate erzielt werden kann.

Eine übliche Vorhergehensweise für die Evaluation eines Softwaresystems ist die top-down Methode. Valutierung und Evaluierung werden direkt vom interagierenden Benutzer durchgeführt. Falls dieser Test nicht erfolgreich ist, würde im nächsten Schritt nach der bottom-up Methode verfahren, wobei während der Verifikation als erstes die Implementierung der operationalen und dann der funktionalen Eigenschaften überprüft werden würde.

Die Evaluation des COGITO Systems basierte auf der top-down Methode, wobei die Interaktion zwischen Mensch und Maschine im Vordergrund stand [Eva02]. Insbesondere erfolgte die Evaluierung und die Valutierung des COGITO Agenten durch Gruppen von Testpersonen, welchen verschiedene Aufgaben, wie z. B. die Suche nach generischen Informationen oder mehr spezifische Produktinformationen, gestellt wurden. Die Evaluation basierte zum Teil auf quantitativen Maßen, wie z. B. die Länge der eingegebenen Sätze, die stereotypische Nutzung der Textausgaben des Agenten, und die Anzahl der „fallback“ Sätze, also Antworten auf nicht interpretierbare Benutzereingaben. Außerdem war die Evaluation qualitativ abhängig von den subjektiven Benutzermeinungen über die Systembedienung sowie über die Ergebnisse einer Informationssuche. Die Evaluation erfolgte durch direkte Befragung und Ausfüllen detaillierter Fragebögen über verschiedene Aspekte, wie z. B. Eindruck, Kontrolle, Effektivität, Navigationsfähigkeit, Lernfähigkeit, Benutzerfreundlichkeit und Verständlichkeit des Agenten.

Vier Gruppen von je acht Personen wurden eingesetzt für die Testsequenzen, unterteilt in zwei Gruppen von Anfängern und zwei Gruppen von erfahrenen Benutzern, damit die Effektivität des COGITO Prototypen für jede dieser Benutzertypen getestet werden konnte.

Damit eine Referenz für die Evaluation des proaktiven COGITO Agenten möglich war, wurde die BOL Website, ausgestattet mit einem marktüblichen Chatterbot, getestet und verglichen. Dieser Agent hatte ein standardmäßiges Gesprächspotential und wurde in die BOL Website integriert, indem einfache Produktverweise als einzige Verbindung zu der Produktdatenbasis implementiert waren, d.h. dieser Agent hatte keine proaktiven Eigenschaften.

Die Analyse des Konversationsprotokolls diente zur Bewertung des Gesprächsverlaufs hinsichtlich der Anzahl korrekter Systemaussagen, „fall-back“ Sätze und proaktiver Systemausgaben. (siehe Abbildung 4). Die Angabe „korrekte Systemaussage“ basierte auf der manuellen Analyse und Interpretation von erfolgreichen Elementen der Mensch-Maschine Interaktion, bestehend aus einem Benutzer-Eingabesatz und dem entsprechenden Agenten-Ausgabesatz. Der COGITO Agent zeigt eine bessere Performanz als der BOL Agent bezüglich der korrekten Ausgabekategorie (61% vs. 47%). Es scheint, als ob der COGITO Agent über eine bessere Suchterkennung verfügt; er hat einen umfangreicheren Wortschatz, unter anderem wegen der Einbindung des Thesaurus. Eine andere Begründung liegt wahrscheinlich in der Fähigkeit proaktiver Aussagen, welche immer dann erfolgen, wenn das System aus Eigeninitiative Vorschläge und Antworten ausweist.

Der BOL Agent stellt selbstverständlich auch Fragen, die allerdings genereller und passiver sind und verwendet nicht die semantische Bedeutung der Benutzereingabe für die Folgenantworten.

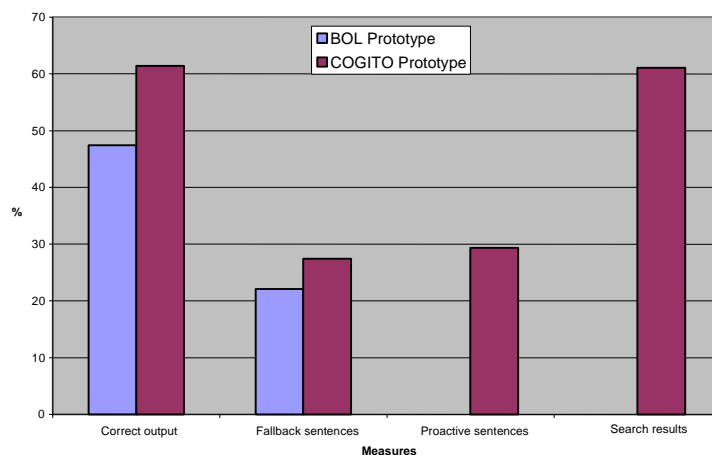


Abbildung 4: Die Ergebnisse aus der Analyse des Gesprächsprotokolls

Aus der Analyse der Suchergebnisse, welche sich aus der automatisch expandierten Anfrage mit Hilfe des Thesaurus ergeben, wird eine Steigerung der Effektivität um insgesamt 61% ersichtlich (siehe letzte Spalte der graphischen

Darstellung in Abbildung 4). Da der konventionelle Chatterbot nur mit den in seinen Regeln vordefinierten „frozen queries“ arbeiten kann, die durch vorher festgelegte Schlüsselbegriffe aktiviert werden, werden hier von den Kunden häufig gebrauchte Anfragen in stereotyper Weise benutzt. Dagegen erweitert der Prompter die als Dialogthemen auftretenden Begriffe unter Verwendung des Thesaurus dynamisch zu einer zur Dialogsituation passenden Anfrage.

6 Ausblick

Der in dieser Arbeit angefertigte Thesaurus befindet sich noch im Prototyp-Stadium. Er lässt sich auf mehrere Arten erweitern. Folgende Veränderungen sind in der Zukunft möglich:

- 1) Erweiterung der Datenbasis
Die Datenbasis in dieser Arbeit beschränkt sich auf Deutsch und die BOL-Produktkategorien. Sie kann aber ohne großen Aufwand auf z.B. englische Webseiten umgestellt werden. Ferner besteht die Möglichkeit, Informationen aus allgemeinen kategoriebasierten Informationsportalen (z.B. Yahoo!²) als Datenbasis für den Thesaurusaufbau hinzuzunehmen, um bessere und allgemeinere Ergebnisse erzielen zu können. Verschiedene Experimente werden durchgeführt, um die Schwellenwerte in den jeweiligen erweiterten Domänen festzustellen. Dabei wird nach universellen Schwellenwerten für alle Domäne gesucht.
- 2) Verbesserung des Thesaurusaufbaus
Es gibt noch einige Verbesserungsmöglichkeiten für den eigentlichen Prozess des Thesaurusaufbaus. Dazu zählen z.B.:
 - Unterscheidung der Informationen in Texttitel und Textkörper, damit die Terme nach ihren Positionen gewichtet werden können
 - Erkennung und Einsetzung von Mehrwortbegriffen
 - Verkleinerung der Größe des „Fensters“, in dem die Terme als relevant betrachtet werden. In dieser Arbeit war das „Fenster“ auf das ganze Dokument gestellt
- 3) Verwendung der „Thematische Spezialisierung“-Beziehung für die automatische Query Expansion
Durch den Einsatz eines Gesprächsprotokolls, welches einen mehrstufigen Dialog zwischen Benutzern und dem System strukturiert aufzeichnet, lassen sich die Benutzerwünsche besser ermitteln. Dadurch wird

² <http://www.yahoo.com>.

nicht nur die „Relevant“-Beziehung, sondern auch die „Thematische Spezialisierung“-Beziehung für die automatische Query Expansion möglich.

- 4) Verwendung des Thesaurus zur interaktiven Query Expansion
Die interaktive Query Expansion wird durch das Einbeziehen der Benutzer bessere Ergebnisse erzielen als die automatische Query Expansion. Dabei muss die Verbindung zwischen dem Thesaurusmodul und dem Agenten verändert werden, um Benutzern verschiedene Expansionsauswahlen anhand der Hierarchie des Thesaurus anzubieten.

7 Literaturverzeichnis

- [Bc92] Belkin, N.J. and Croft, W.B.: Information Filtering and Information Retrieval: Two Sides of the Same Coin? *Communications of the ACM*, 35(12), pages 29-38, December 1992.
- [Bco00] Beschreibung des Projekts COGITO
<http://www.darmstadt.gmd.de/~cogito/Description.htm>.
- [Byf93] Frakes, W.B. and Bareza-Yates, R.: *Information Retrieval Datastructures & Algorithms*, Prentice-Hall, 1993.
- [Byrn99] Bareza-Yates, R. and Ribeiro-Neto, B.: *Modern Information Retrieval*, Addison-Wesley, 1999.
- [CrY92] Crouch, C.J., Yong, B., Experiments in automatic statistical thesaurus construction, SIGIR'92, 15th Int. ACM/SIGIR Conf. on R&D in Information Retrieval.
- [Eft96] Efthimiadis, E. N.: Query Expansion, *Annual Review of Information Systems and Technology (ARIST)*, v31, pp 121-187, 1996.).
- [Ekm92] Ekmekcioglu, F.C., Robertson, A.M., Willett, P.: Effectiveness of query expansion in ranked output document retrieval systems, *J. of Information Science*. 1992
- [Elna94] Elmasri, R., Navathe, S.: *Fundamentals of Database Systems*. Redwood City: The Benjamin/Cummings Publishing Co., Inc., 1994.
- [Eva02] Project Deliverable report, D7.2 Evaluation of the COGITO system.
- [Ferb97] Ferber, R.: *Data Mining und Information Retrieval*. Vorlesungsskript 1997
<http://www.muenster.de/~ferber/>.
- [From01] Frommholz, I.: *Automatische Kategorisierung von Web-Dokumenten*. Diplomarbeit, 2001
<http://ls6-www.cs.uni-dortmund.de/bib/fulltext/world/Frommholz:01.ps.gz>.
- [Fuhr96] Fuhr, N.: *Information Retrieval*. Skriptum zur Vorlesung. Technical Report. Universität Dortmund, Fachbereich Informatik. 1996
<http://ls6-www.cs.uni-dortmund.de/ir/teaching/courses/ir/>.
- [Göv99] Gövert, N., Lalmas, M. and Fuhr, N.: A probabilistic description-oriented approach for categorizing web documents. In Susan Gauch and Il-Yeol Soong, editors, *Proceedings of the Eighth International Conference on Information and Knowledge Management*, pages 475–482, New York, 1999.

- [Jicr94] Jing, Y. F. and Croft, W. B.: An Association Thesaurus for Information Retrieval, UMass Technical Report 9417.
- [Klas00] Klas, P. and Fuhr, N.: A new effective approach for categorizing web documents. In Proceedings of the 22th BCS-IRSG Colloquium on IR Research, 2000.
- [Kow97] Kowalski, G.: Information Retrieval Systems: Theory and Implementation. Kluwer Academic Publishers: Boston/Dordrecht/London, 1997.
- [Lath01] L'Abbate, M. & Thiel, U.: Chatterbots and intelligent Information Search. Proceedings of ECIR 2001, Darmstadt, Germany 2001.
- [Mink72] Minker, J., Wilson, G.A. and Zimmermann, B. H.: An evaluation of query expansion by the addition of clustered terms for a document retrieval system, Information Storage and Retrieval, 8, 329-348 (1972).
- [Panyr86] Panyr, Jiri: Automatische Klassifikation und Information Retrieval: Anwendung und Entwicklung komplexer Verfahren in Information-Retrieval-Systemen und ihre Evaluierung. (Sprache und Information, Bd. 12). Tübingen: Niemeyer, 1986, 416 S.
- [Qiu93] Qiu, Y., Frei, H.P.: Concept based query expansion. In Proceedings of ACM SIGIR International Conference on Research and Development in Informaiton Retrieval.
- [Rakr99] Robinson, T.; Abberley, D.; Kirby, D. and Renals, S.: Recognition Indexing and Retrieval of British Broadcast News with the THISL System. Proc. Eurospeech-99 Budapest, Hungary, 1067-1070 (1999).
- [Rij79] Rijsbergen, C. J. Van: Information Retrieval. Butterworth, London, second edition, 1979.
- [Rocchio71] Rocchio, J.Y. : Relevance Feedback in Information Retrieval. The SMART Retrieval System. Englewood Cliff, N.J.: Prentice Hall, 313-323.
- [Salton73] Salton, G., Comment on "an evaluation of query expansion by the addition of clustered terms for a document retrieval system". Computing Reviews, 14, 232 (1973).
- [Same83] Salton, G. and McGill M.: Information Retrieval - Grundlegendes für Informati-onswissenschaftler. 1983.
- [Spa71] Sparck-Jones, K.: Automatic Keyword Classification for Information Retrieval. Butterworth, London 1971.
- [Stef96] Steffens, U.: Integration von Information Retrieval Funktionalität in eine offene persistente Programmierumgebung. Informatik Mitteilung FBI-HH-M-257/96, Fachbereich Informatik, Universität Hamburg, Germany, April 1996.
- [Thl00] Thiel, U.; Stein, A.; Semeraro, G.; Abbatista, F.; De Candia, L.; Fanizzi, N.; Candela, V.; Lops, P. & Valente, A.: COGITO - E-Commerce with Guiding Agents Based on Personalized Interaction Tools. In Proceedings of the AICA Annual Conference, Taormina, Italy, September 27-30, 2000.
- [Xucr96] Xu, J. and Croft, W. B.: Query expansion using local and global document analysis, in Proc. ACM SIGIR, 1996.



Deutsche Suchmaschinen im Vergleich: AltaVista.de, Fireball.de, Google.de und Lycos.de

Joachim Griesbaum¹, Marc Rittberger², Bernard Bekavac¹

¹Universität Konstanz
Informationswissenschaft

Fach D 87

D-78457 Konstanz

{Joachim.Griesbaum,

Bernard.Bekavac}@uni-konstanz.de

²Heinrich-Heine-Universität
Düsseldorf

Institut für Sprache und Information

D-40225 Düsseldorf

Marc@Rittberger.de

Zusammenfassung

Die vier Suchmaschinen AltaVista.de, Fireball.de, Google.de und Lycos.de werden einem Retrievaltest unterzogen, um ihre Eignung für den deutschsprachigen Suchraum zu betrachten. Die Evaluierung erfolgt mit 28 Studierenden und Mitarbeitern der Informationswissenschaft und insgesamt 56 Suchfragen im Januar 2002. Es zeigen sich deutliche Vorteile für Google.de gegenüber den anderen Suchmaschinen. Die sichtbaren Vorteile von Lycos.de können den statistischen Überprüfungen nicht standhalten, so dass bei den anderen drei Suchmaschinen von einer gleich hohen Retrievalleistung ausgegangen werden muss.

1 Einleitung

Transparenz der Ergebnisse und Qualität des Rankings bei Suchmaschinen beschäftigen die Öffentlichkeit und die Wissenschaft immer mehr. Die Aktualität der Suchergebnisse, besser des Index, die möglichst vollständige Erfassung der Inhalte des Internets, aber auch Stichworte wie Spamming¹, Cloaking² oder Doorway-Pages³ spielen eine große Rolle in den Diskussionen.⁴

¹ Spamming bezeichnet neben dem Überfluten mit unerwünschten Informationseinheiten durch Werbemails oder Marketingmaßnahmen in News-Groups im Umfeld von Suchmaschinen eher unlautere Verfahren, die zu einem hohen Ranking insbesondere bei nur näherungsweise relevanten Seiten führen.

² Cloaking ist das IP-abhängige Zurverfügungstellen unterschiedlicher Versionen ein und derselben Webseite.



Möglichst hohe Aktualität, große Vollständigkeit und genaue Ergebnisse sind zwar auch im Interesse der Suchmaschinen, um eine hohe Kundenzufriedenheit zu erreichen, aber durch die Größe und die Vielfalt der Daten, die auf dem Internet angeboten werden, nur schwer zu erreichen.

Geht es um die Bewertung von Suchmaschinen, so spielt die Diskussion um die Positionierung von Ergebnissen im Ranking und damit natürlich die Relevanz eine herausragende Rolle. Das hat in den letzten Jahren dazu geführt, dass in fast jeder Internet- oder Computer-Zeitschrift Suchmaschinentests veröffentlicht werden (z.B. [Bager & Schulzki-Haddouti 2001; o.V. 2001]). In solchen Tests werden die jeweilige Testanordnung und Testdurchführung jedoch nur selten umfassend und transparent dargestellt,⁵ womit unklar bleibt, ob und inwieweit die erzielten Testresultate neutral und „objektiv“ und damit für Suchdienstennutzer tatsächlich hilfreich sind.⁶

Zur Nutzung von Suchmaschinen und den daraus resultierenden Ergebnissen gibt es mit Bezug auf die internationale Ausrichtung von Suchmaschinen zahlreiche Untersuchungen (z.B. [Bar-Ilan 2002; Su & Chen 1999; Wang et al. 1999; Gordon & Pathak 1999; Leighton & Srivastava 1999; Hawking et al. 2001; Ford et al. 2001; Mettrop 2001; Dennis et al. 2002]).

Neben den gerne genutzten internationalen Suchdiensten wird die nationale Ausrichtung von Suchmaschinen allerdings immer wichtiger. Die Zuwachsraten in Sprachräumen oder auf Nationen bezogen, die nicht dem Englischen angehören, sind deutlich höher als die Veränderungen, die sich in den durch das Englische dominierten Informationsmärkten ergeben. Die Nationalisierungsstrategien der Anbieter [Dresel et al. 2001], die damit den Anforderungen der Informationsmärkte folgen, bestätigen diesen Eindruck. Bei einer Befragung von 43 Studierenden der Informationswissenschaft in Düsseldorf,

³ Doorway-Pages sind speziell für Suchroboter optimierte Seiten, die insbesondere relevante Texte, Indexierung etc. enthalten.

⁴ Mit Spamming, Cloaking oder Doorway-Pages werden Verfahren beschrieben, mit denen Anbieter von Inhalten auf dem World Wide Web versuchen, ihre Seiten bei den Suchmaschinen möglichst gut zu platzieren, ohne dass die Inhalte, die über die genannten Verfahren suggeriert werden, wirklich vorhanden sind.

⁵ Siehe z.B. Newsmeldung vom 13.05.02 bei Suchfibel.de „Seltsamer PC Pro Suchmaschinentest“.

URL <http://www.suchfibel.de/news/SeltsamerPCProSuchmaschi.htm> [17.05.02].

⁶ Eine vergleichende Gegenüberstellung solcher Tests findet sich auf der Webseite von Klaus Patzwald.

URL <http://www.at-web.de/Informationen/suchmaschinen-test.htm> [24.04.02].

Konstanz und Chur wurde die Wichtigkeit dieser Einschätzung bestätigt, da die Mehrheit der Studierenden der Einschränkung der Suchergebnisse auf nationale, mehr noch auf Sprachraumgrenzen eine hohe Bedeutung gab.

Umso verwunderlicher ist es daher, dass zumindest für den deutschsprachigen Informationsmarkt kaum wissenschaftliche Untersuchungen und Ergebnisse über die Leistungsfähigkeit der auf diesen Sprachraum fokussierten Suchmaschinen zur Verfügung stehen.⁷ Der Frage nachzugehen, wie gut die deutschsprachigen Suchmaschinen sind, scheint daher durchaus eine Lücke zu schließen.

Dieser allgemeine Mangel an aktuellen Studien zum Vergleich von deutschen Suchmaschinen wird im Januar 2002 in Zusammenarbeit der informationswissenschaftlichen Forschungsgruppen in Konstanz und Düsseldorf zum Anlass genommen, einen Retrievaltest der Suchmaschinen Altavista.de, Fireball.de, Google.de und Lycos.de durchzuführen. Der Retrievaltest verfolgt zwei Ziele, erstens die Retrievaleffektivität der untersuchten Suchmaschinen zu ermitteln und zweitens das in [Griesbaum 2000] konzipierte Evaluationsverfahren, auf dem dieser Retrievaltest aufbaut, weiter zu entwickeln und damit für die systematische Durchführung von Retrievaltests mit Suchmaschinen weitere Erfahrungen zu gewinnen.

2 Vorgehensweise und Zielsetzung

Die vorliegende Evaluation beruht methodisch primär auf der von [Tague-Sutcliffe 1992], S.467 vorgeschlagenen grundlegenden Vorgehensweise zur Entwicklung eines Retrievaltests.⁸ Die Testart folgt dabei dem Cranfield Paradigma [Ellis 1992] und ist grundsätzlich bemüht, z.B. in Anlehnung an TREC, etablierte Evaluationsstandards einzuhalten.⁹

Die Vorgehensweise nach [Tague-Sutcliffe 1992] unterteilt einen Retrievaltest in zehn Teilbereiche, welche sequentiell abzuarbeiten sind:¹⁰

1. Testen oder nicht testen (Need for testing) - Motivation des Retrievaltests

⁷ Ausnahmen sind [Dresel et al. 2001; Stock & Stock 2000; Wolff 2000].

⁸ Dieser „Leitfaden“ löst zwar nicht die elementare Problematik jedes Retrievaltests an sich, die der adäquaten Ausgestaltung der quantitativen und qualitativen Ausgestaltung der Testparameter - kann sie nicht lösen - bietet aber in den vorgeschlagenen zehn sequentiell zu durchlaufenden Schritten grundlegende Hinweise und Entscheidungshilfen, welche die Validität, Reliabilität und Effizienz des Testsettings absichern sollen.

⁹ Beispielsweise bei der Anzahl der Suchanfragen.

¹⁰ In Klammern stehen die englischen Originalbezeichnungen für die einzelnen Stufen eines Retrievaltests.

2. Testart (Type of test) - Bestimmung des grundsätzlichen Testverfahrens
3. Variablendefinition und -zuordnung (Definition of variables)
4. Verwendetes Informationssystem (Database development) - Ausgewählte Suchmaschinen
5. Erschließung der Informationsbedürfnisse und Suchanfragen (Finding queries)
6. Durchführung der Suchanfragen (Retrieval software)
7. Testanordnung (Experimental design)
8. Datenerfassung (Data collection)
9. Datenauswertung (Data analysis)
10. Ergebnispräsentation (Presenting Results)

Steht das Testsetting, wird durch einen Pretest überprüft, ob das Evaluationssetting tatsächlich adäquat ausgestaltet ist. In Abhängigkeit der Resultate des Pretests wird die Testanordnung gegebenenfalls angepasst bis dieses Ziel sichergestellt scheint. Sind diese Voraussetzungen erfüllt, werden die Tests durchgeführt, die Daten ausgewertet und analysiert.

Um einerseits die pragmatische Handlungsrelevanz der Ergebnisse im realen Nutzungskontext der untersuchten Suchdienste einschätzen zu können und andererseits Probleme bei der Durchführung der Untersuchung aufzuzeigen und – idealerweise – Optimierungspotenziale für künftige Retrievaltests zu erschließen, werden abschließend sowohl die Ergebnisse als auch die Untersuchung selbst kritisch hinterfragt.

Innerhalb dieser „theoretisch-normativen“ Rahmenstruktur wird bei der konkreten Ausgestaltung der einzelnen Testparameter versucht, webspezifische Eigenheiten des Information Retrieval hinsichtlich Datenbestand, Hypertextstrukturen und Nutzerverhalten möglichst realitätsnah nachzubilden. [Gordon & Pathak 1999] nennen sieben Kriterien zur Evaluierung von Suchmaschinen im Web als essentiell, die von [Hawking et al. 2001] kritisch hinterfragt wurden und zu Recht auf fünf Kriterien reduziert werden:

1. Reale Informationsbedürfnisse von Nutzern sollen abgebildet werden.
2. Bei der Einbindung von Informationsvermittlern soll das originäre Informationsbedürfnis sorgfältig mitgeteilt werden.
3. Es soll eine große Anzahl von Suchfragen genutzt werden.
4. Die wichtigsten Suchmaschinen sollen involviert sein.
5. Die Untersuchung soll gut und sorgfältig aufgebaut und durchgeführt werden.

Diese fünf Kriterien, die bei einer Suchmaschinenevaluierung erfüllt sein sollen, werden von [Hawking et al. 2001] um das Kriterium, dass die Frageformulierung das Informationsbedürfnis möglichst gut treffen soll, ergänzt.

Wir werden im Rahmen der Systematik von [Tague-Sutcliffe 1992] die einzelnen Schritte unseres Retrievaltests im nächsten Abschnitt vorstellen. Ausreichend diskutiert erscheint uns dabei schon der erste Punkt der Motivation des Retrievaltests. Auch werden wir die Ergebnispräsentation ausführlicher in einem eigenen Abschnitt darstellen.

3 Durchführung des Tests

3.1 Teststart – Bestimmung des grundsätzlichen Testverfahrens

Das gewählte Testverfahren ist ein Test anhand einer Testkollektion. Die Informationsbedürfnisse, Suchanfragen, Bewertungsmaße und Bewertungskriterien werden den Testpersonen (Juroren) somit extern von den Untersuchenden vorgegeben. Dieses dem Cranfield-Paradigma folgende, an die TREC ad hoc und Web Track Tasks [Voorhees & Harman 2001] angelehnte, insgesamt eher laborhaft ausgestaltete Testverfahren bildet die reale Nutzungssituation bei Suchdiensten zwar nur unvollkommen nach, ermöglicht aber gerade durch die „standardisierte“ Testanordnung eine Abstraktion von individuell wirkenden, schlecht zu kontrollierenden Einflussfaktoren und sichert damit die Vergleichbarkeit der einzelnen Testergebnisse [Griesbaum 2000], S.26, 58f.

3.2 Variablendefinition und –zuordnung

Die Variablen Relevanzeinstufung, Dokumentdarstellung, Bewertungsmaße, Suchanfragen und Informationsbedürfnisse, Testpersonen, Dokumentraum, und Relevanzeinstufung des Retrievaltests sowie die Handhabung der Variablen werden festgelegt.

Relevanzeinstufung

Als zentrales Konzept zur Beurteilung von Retrievalsystemen gilt die Relevanz der gelieferten Treffer [Robertson 1981], S.14. Dieses trefferbasierte Bewertungsmaß bildet trotz der Problematik der personengebundenen Subjektivität jedes Relevanzurteils, mangels Alternativen, auch in diesem Test die Grundlage der Bewertungsmaße [Warner 2000], S.77.

Die Relevanzeinstufung der Treffer wird von Testpersonen, die als Juroren fungieren, vorgenommen. Dies gewährleistet zumindest, dass die Bewertungen von Vorlieben bzw. Abneigungen der Untersuchenden unbeeinflusst bleiben [Robertson 1981], S.17. Um dies auch bei den Juroren selbst sicherzustellen, wird die Herkunft der Treffer unkenntlich gemacht. Um die Eindeutigkeit der Relevanzurteile abzusichern, wird jedes Dokument nur von einem Juror bewertet.

Durch die Hypertextstruktur des Web ist es denkbar, dass ein an sich irrelevantes Dokument über Verknüpfungen den direkten Zugriff auf relevante Seiten ermöglicht. Aus diesem Grund wird von einer binären Relevanzeinstufung in „relevant“ und „nicht relevant“ zunächst abgesehen und als dritte Bewertungsmöglichkeit „verweist auf relevante Seite(n)“ hinzugefügt.¹¹ Weil solche als „verweist auf relevante Seite(n)“ bewerteten Treffer letztendlich hilfreich sind, um Informationsbedürfnisse zu befriedigen, werden sie bei der Auswertung zu den relevanten Treffern gezählt. Dies reflektiert auch das tatsächliche Verhalten der Suchdienstenutzer, bei denen während der Nutzung der Suchdienste die Browsing-Perioden überwiegen [Körber 2000], S.41.

Problematisch ist, dass bei der Relevanzbewertung von einer Einzelbetrachtung der Dokumente ausgegangen wird. Damit wird beispielsweise ein gleichbleibender Wissensstand des Jurors impliziert. Die sogenannten „Grenzfälle der Relevanz“, vgl. [Griesbaum 2000], S.64, werden damit bis auf die Berücksichtigung von Dubletten, die nur beim ersten Auftreten als relevant gewertet werden können, nicht berücksichtigt.

Dokumentdarstellung

Die Relevanzeinstufung wird anhand der Ergebnisseiten selbst vorgenommen. Die Juroren haben eine Liste von Links im Browser vor sich, die auf die relevanten Seiten der von den Suchmaschinen gefundenen Treffer verweisen.

Bewertungsmaße

Recall und Precision sind die Standardwerte zur Effektivitätsmessung von Retrievalsystemen [Lesk 1995]. Auf Recall wird in dieser Untersuchung verzichtet, da er zum einen im Web nicht oder nur unzureichend bestimmt werden kann [Oppenheim et al. 2000], S.190, und zum anderen der vollständige Nachweis aller relevanten Dokumente, gerade für Nutzer im Web häufig nur

¹¹ Dies reflektiert auch das tatsächliche Verhalten der Suchdienstenutzer, bei denen während der Nutzung der Suchdienste die Browsing-Perioden überwiegen [Körber 2000], S.41.

von geringem Interesse ist, da sie in der überwiegenden Zahl der Fälle nur die ersten zwei, selten die ersten drei Ergebnisseiten¹² der Suchmaschinen sicherten.¹³ Dieser Retrievaltest beschränkt sich deshalb auf die Effektivitätsbeurteilung anhand der Relevanzbeurteilung der ersten 20 Treffer, der sogenannten Top20 Precision.

Bei der Auswertung werden sowohl die Mikro- als auch die Makromethode verwendet. Bei der Makromethode werden zunächst die einzelnen Suchanfragen als Grundeinheit betrachtet, d.h. zuerst werden die Precisionwerte pro Suchanfrage berechnet und dann die Werte der Suchanfragen gemittelt, damit fließt jede Suchanfrage gleichgewichtig in die Bewertung ein. Bei der Mikromethode hingegen werden die Dokumente als Grundeinheit genommen, sie berechnet, unabhängig von der Trefferanzahl der einzelnen Suchanfragen, das Verhältnis der relevanten Treffer zu allen Treffern, dadurch fließt jedes Dokument gleichgewichtig in die Bewertung mit ein [Womser-Hacker 1989], S.66f .

Suchanfragen und Informationsbedürfnisse

Die Informationsbedürfnisse und Suchanfragen sind das zentrale Element jedes Retrievaltests. Die Thematik, Komplexität, Spezifität der Fragestellungen und die Art der Suchanfragenformulierung spezifizieren die inhaltliche Ausprägung der Untersuchung und determinieren so als wichtigste Inputfaktoren bei der Testdurchführung unmittelbar die Quantität und Qualität der Treffer der Suchmaschinen.¹⁴

Die Fragestellungen werden in dieser Untersuchung nicht den „typischen“ Fragestellungen im Web nachgebildet, sondern als normative Ausprägung des Retrievaltests festgelegt.

Der Themenbereich wird zunächst durch Abgrenzung unerwünschter Bereiche negativ spezifiziert. Explizit ausgeschlossen werden Fragestellungen, die

¹² Gemeint sind hier die Trefferlisten.

¹³ Laut AltaVista.com benutzen sogar weniger als 10% die zweite Ergebnisseite, vgl. [Körber 2000], S.33 und [Jansen et al. 2000].

¹⁴ Zur Veranschaulichung folgendes Beispiel:

Zur Suchanfrage „hypothesengenerierende Untersuchungsverfahren“ in Phrasenform liefern Google.de, Lycos.de, Fireball.de und bei Altavista.de keinen Treffer. Bei der Suchanfrage „Star Wars Klonkriege Premiere Termin“ ohne Phrasenform referenziert Google.de 4 Treffer, Lycos.de 3 Treffer, Fireball.de 2 Treffer und Altavista.de 1 Treffer, die Suchanfrage „Autokauf“ liefert bei Google.de 93.800 Treffer, bei Lycos.de 23888 Treffer, bei Fireball.de 14996 Treffer und bei Altavista.de 17765 Treffer.

Anfragen durchgeführt am 30.04.02.

auf materiellen Bedürfnissen beruhen und auf Produkt- oder Dienstleistungsangebote kommerzieller Anbieter zielen. Ebenso ausgegrenzt werden Fragestellungen, die direkt auf Freizeitaktivitäten oder Hobbies schließen lassen. Fragestellungen aus dem Porno/Erotik-Bereich werden ebenfalls ausgeklammert.

Die Anzahl der Suchanfragen wird mit 50, entsprechend der Standardvorgabe von TREC, hinreichend hoch gewählt, um verallgemeinerungsfähige Testergebnisse zu erhalten. [Buckley & Voorhees 2000], S.33. Damit ein Puffer zur Verfügung steht, werden 60 Suchanfragen vorbereitet.

Die Formulierung der Suchanfragen wird den typischen Nutzergewohnheiten angelehnt, die überwiegend Suchwörter ohne Operatoren oder Klammerung eingeben [Jansen et al. 2000].

Testpersonen

Zur Verfügung stehen die Teilnehmer der Information Retrieval Vorlesungen der Informationswissenschaft an den Universitäten in Düsseldorf und Konstanz. Alle Teilnehmer sind fortgeschrittene Studierende, die sowohl im Umgang mit dem Web als auch mit Suchmaschinen über viel Erfahrung verfügen.

Dokumentraum Internet

Das Internet ist im Rahmen dieses Retrievaltests als nicht zu beeinflussender Dokumentraum einzuordnen. Die Datenbestände des Internet sind sehr heterogen und dynamisch, bieten aber den Suchdiensten und ihren Indexierungskomponenten dieselben Bedingungen.

Relevanzeinstufung der Treffer

Suchdienste liefern in der Regel eine nach maschinell berechneter Relevanz sortierte Liste von Links (Treffern), die den Suchdienstbenutzer direkt zu den Ergebnisseiten führen. Die Listen werden im Testumfeld so maskiert, dass die Juroren nicht erkennen können, von welcher Suchmaschine welche Treffer stammen.

3.3 Ausgewählte Suchmaschinen

Es werden vier Suchmaschinen miteinander verglichen. Die Auswahlkriterien sind Reichweite und Nutzungsgrad. Die Analyse einschlägiger statistischer Quellen¹⁵ ergibt folgende Auswahl: Altavista.de, Fireball.de, Google.de und

¹⁵ Vgl. Jupiter MMXI, European Search Engine Ratings.

Lycos.de. Mitausschlaggebend für die Auswahl war die Untersuchung der Stiftung Warentest [o.V. 2001], die die genannten vier Suchmaschinen als einzige mit einem guten (Google.de) oder zumindest befriedigendem Abschneiden bei der Bewertung der Suchergebnisse versehen hat.

3.4 Erschließung der Informationsbedürfnisse und Suchanfragen

Das grundlegende Problem der „Repräsentativität“ von Suchanfragen und Informationsbedürfnissen kann im Rahmen dieses Tests nicht gelöst werden. Um aber bei der Auswahl eine thematische oder sonstige wie auch immer geartete Verengung zu vermeiden, werden bei der Erschließung der Suchanfragen verschiedene Quellen berücksichtigt.

<i>Quelle</i>	<i>Anzahl Suchanfragen</i>
Cross Language Track (Trec 7)	7
Cross Language Track (Trec 8)	5
GIRT (Trec 8)	17
Web Track (Trec 8)	4
Ask Jeeves	27

Tabelle 1: Die Quellen der Suchfragen

<i>Anzahl offene Fragestellungen</i>	<i>Anzahl geschlossene Fragestellungen</i>
11	49

Tabelle 2: Anzahl offener und geschlossener Suchfragen

<i>Anzahl Suchbegriffe</i>	<i>Anzahl Suchfragen</i>
1	1
2	30
3	18
4	8
5	1
6	1

Tabelle 3: Anzahl der Suchbegriffe

Deshalb werden zum einen Informationsbedürfnisse und Fragestellungen aus verschiedenen Teilkollektionen von Trec verwendet (Cross Language Topics, GIRT, Web Track), zum anderen aber auch neue „real existierende“ Fragestellungen durch Auswertung von Anfragen im Web erschlossen. Dabei wird, in Anlehnung an die Fragestellungen von Web Track, der Fragenprotokoll-dienst von Ask Jeeves ausgewählt.¹⁶

Tabelle 1 zeigt die Herkunft der Suchfragen an, Tabelle 2 die Anzahl der offenen und geschlossenen Suchfragen und Tabelle 3 die Anzahl der Suchbeg-

URL <http://www.searchenginewatch.com/reports/mmxi-europe.html> [05.04.02],

URL <http://www.webhits.de/deutsch/index.shtml?/deutsch/webstats.html> [03.04.02].

¹⁶ <http://www.askjeeves.com/docs/peek/> [05.05.02].

riffe in den einzelnen Fragen. Demnach gibt es bspw. acht Suchfragen, die mit vier Suchbegriffen operieren.

3.5 Durchführung der Suchanfragen

Das Ziel, die Validität, Reliabilität und Effizienz des Retrievaltests im Vergleich zur vorangegangenen Untersuchung [Griesbaum 2000] zu erhöhen, soll durch eine Automatisierung bei der Durchführung der Suchanfragen erreicht werden.

In [Griesbaum 2000] wurden die Suchanfragen händisch durchgeführt. Dieser Prozess, insbesondere die zur Verfügungstellung der Suchergebnisse, war sehr zeitaufwändig - Dauer rund zwei Wochen - und die originalgetreue Nachbildung einiger, insbesondere dynamischer Ergebnisseiten gelang nur näherungsweise. Diese Probleme werden in der vorliegenden Evaluierung konzeptionell durch ein Abfrageskript gelöst, welches es gestattet, die Suchanfragen erst unmittelbar vor der Relevanzbeurteilung der Dokumente durchzuführen. Damit sollen die zeitlichen Verzerrungen zwischen Anfrage und Bewertungszeitpunkt minimiert werden, wodurch die Validität der Untersuchung erhöht wird. Da damit zugleich die fehlerbehaftete lokale Spiegelung der Ergebnisseiten überflüssig wird, werden weniger Ressourcen benötigt, was wiederum zu einer Effizienzsteigerung führt.

3.6 Testanordnung

Aufgrund der Menge der Suchanfragen und der Anzahl und Verteilung der Juroren in Konstanz und Düsseldorf wird entschieden, den Test innerhalb von zwei Tagen durchzuführen. Die Auswirkungen der zeitlichen Unterschiede werden als hinnehmbar betrachtet, solange keine der Suchmaschinen innerhalb des Testzeitraums größere Modifikationen vornimmt.¹⁷ Die Juroren haben jeweils insgesamt zwei Suchanfragen zu beurteilen. Der zeitliche Aufwand wird pro Termin auf rund eine Stunde geschätzt, wobei eine breite Streuung der benötigten Zeit von rund einer halben bis zu zwei Stunden erwartet wird.

Im Testablauf ist es die Aufgabe der Juroren, nacheinander die Treffer 1-20 der vier Suchdienste zu beurteilen. Diese Anordnung entspricht am ehesten einem „Repeated Measures Design“ [Tague-Sutcliffe 1992]. Die dabei auftre-

¹⁷ Etwa vergleichbar mit dem Relaunch von Fireball.de am 04.04.2002.
URL <http://www.tecchannel.de/news/20020404/thema20020404-7194.html> [05.05.02].

tenden Lern- und Ermüdungseffekte sollen durch die Variation der Reihenfolge der Trefferlisten der Suchmaschinen kompensiert werden.

3.7 Datenerfassung

Es werden die Bewertungen und Anmerkungen zu den Ergebnisseiten aufgenommen und die persönlichen Angaben über die Juroren hinsichtlich Alter, Geschlecht, Qualifikation und Gemütsbefinden erfasst.

3.8 Datenauswertung

Zunächst werden die Relevanzeinstufungen „Relevant“ und „Verweist auf relevante Seite(n)“ zusammengeführt und als relevant bewertet, um zu einem binären Werturteil zu kommen, welches für die Berechnung der Retrievalmaße notwendig ist.

Die Retrievaleffektivität der Suchmaschinen wird konkret anhand der Maßzahl der Top20 Precision überprüft. Dabei interessieren primär zwei Ergebnissichten. Zum einen die Makroprecision, die zeigt, wie effektiv die Suchmaschinen die einzelnen Suchanfragen beantworten, und zum anderen die Mikroprecision, die aufzeigt, welche Effektivität die Suchdienste über alle Suchanfragen hinweg bei den Cut-off Werten 1-20 erreichen

Zur Ermittlung und Absicherung der statistischen Validität der Ergebnisse werden diese auf Signifikanz überprüft. Erst dadurch ist es möglich zu entscheiden, ob die ermittelten Unterschiede hinreichend sind, um verallgemeinernde Schlüsse zu ziehen.

4 Pretest

Der Pretest wird am 04.01.02 durchgeführt. Bei der Durchführung der Suchanfragen gibt es Schwierigkeiten. Einige Suchmaschinen, vor allem Google, liefern aus nicht nachvollziehbaren Gründen bei einigen Anfragen zunächst keine Treffer bzw. bei wiederholter Anfrage unterschiedliche Trefferzahlen.

Deshalb werden beim realen Test die Suchanfragen bereits einen Tag vor dem ersten Termin, an dem die Relevanzbeurteilung durch die Juroren vorgenommen wird, durchgeführt und nicht jeweils unmittelbar vor den Terminen. Diese Entscheidung wirft Probleme auf. Denn die Trefferlisten werden zwar zum gleichen Zeitpunkt generiert, die Ergebnisseiten sind allerdings zum

Zeitpunkt der Relevanzbeurteilung schon mindestens einen Tag „veraltet“. Damit besteht die Gefahr bzw. die hohe Wahrscheinlichkeit, dass sich zumindest bei einigen Ergebnisseiten innerhalb dieses Zeitraums die Inhalte verändern oder gar nicht mehr auf die Seiten zugegriffen werden kann.¹⁸

Im Gegensatz zur Durchführung der Suchanfragen verläuft die Relevanzbeurteilung unproblematisch.¹⁹ Der Test dauert eine Stunde und der Testjuror selbst bewertet die benötigte Zeitdauer und die persönliche Belastung als unproblematisch, die Aufgabenstellung als intuitiv verständlich und den Ablauf als klar gekennzeichnet.

Um zu überprüfen, ob das Verschleiern der Herkunft der Treffer trotz der Tatsache, dass ausschließlich Google PDF-Files als Treffer liefert und Katalogtreffer mit dem String „directory.google.com“ kennzeichnet, auch bei Experten funktioniert, wird die Pretestsuchanfrage besonders ausgestaltet. Und zwar derart, dass die Trefferliste von Google bei 8 Treffern PDF-Dokumente als Ergebnisseiten liefert. Der Juror antwortet nach dem Pretest auf die Frage, ob er die Trefferlisten einzelnen Suchmaschinen zuordnen konnte, mit nein. Dies ist ein starker Hinweis darauf, dass die Präsenz von PDF-Dokumenten in den Trefferlisten nicht dazu führt, dass die Juroren Google als Quelle dieser Treffer assoziieren.

5 Testdurchführung

Die Suchanfragen werden am 14.01. in Konstanz durchgeführt. Die Probleme unterschiedlicher Treffermengen treten bei einigen Suchanfragen wieder auf und werden durch wiederholte Durchführung der entsprechenden Anfragen kompensiert. Die Relevanzbeurteilungen werden am 15.01. und 16.01. durchgeführt. Die Testpersonen sind die Teilnehmer der Retrievalkurse in Konstanz und Düsseldorf. In Konstanz werden zusätzlich noch einige Mitarbeiter im Fachbereich rekrutiert. Insgesamt können die Suchtreffer von 56 der 60 vorbereiteten Suchanfragen auf Relevanz bewertet werden.

Die Relevanzbeurteilungen verlaufen unproblematisch. Hinsichtlich der getesteten Suchmaschinen gibt kein Juror einen Hinweis dahingehend, dass er die Trefferlisten einzelnen Suchdiensten zuordnen kann. Vielmehr werden die

¹⁸ Beispielsweise bei URL not found (404) Seiten, die gelöscht oder verschoben wurden.

¹⁹ Die speziell für den Pretest vorbereitete Suchanfrage findet sich unter http://www.infwiss.uni-konstanz.de/CURR/winter0102/IR/uebungen_ir_ws0102/suchanfragen/63_ghmb_o.html [07.05.02].

Untersuchenden von den Juroren nach dem Test oft explizit gebeten, doch mitzuteilen, welche Trefferliste welcher Suchmaschine zuzuordnen ist.

6 Ergebnisanalyse

6.1 Übersicht Anzahl relevanter Treffer

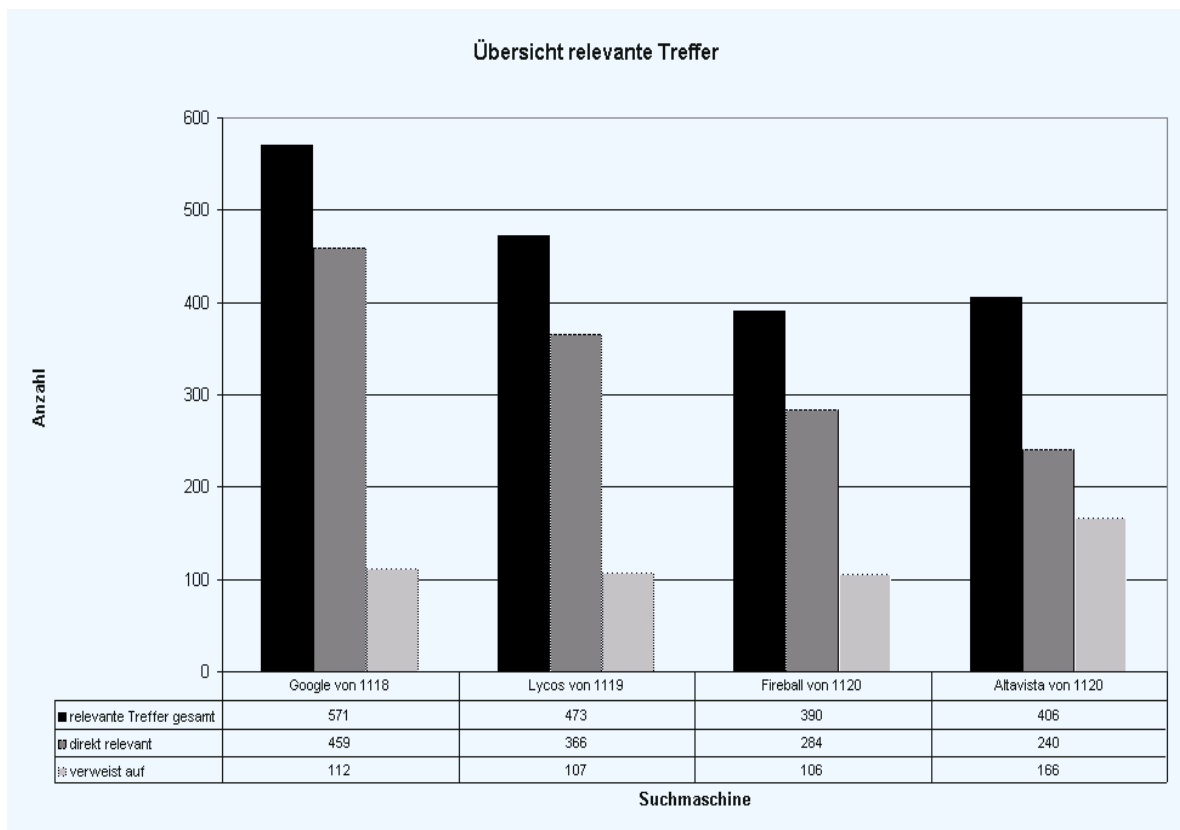


Abbildung 1: Anzahl der relevanten Treffer im Verhältnis zu allen Treffern für die Suchmaschinen google.de, lycos.de, fireball.de, altavista.de

Die Übersicht über die relevanten Treffer in Abbildung 1 zeigt an, wie viele Treffer von der pro Suchmaschine maximal erreichbaren Trefferzahl von 1120 Treffern als relevant beurteilt wurden.²⁰ Während Google über 50% re-

²⁰ Das bedeutet nicht, dass jede Suchmaschine so viele Treffer zurücklieferte, tatsächlich ist dies nur bei Altavista (wahrscheinlich aufgrund der disjunktiven Verknüpfung der Suchbegriffe) der Fall. Google referenziert zu allen Suchanfragen auf den ersten 20 Rangplätzen 1084 Treffer, Lycos 1082 und Fireball 1024 Treffer. Würde man also die Precision auf die tatsächliche Trefferanzahl innerhalb der ersten zwanzig Rangplätze beziehen und nicht auf die ersten 20 Rangplätze, so würde diese bei den letztgenannten Maschinen leicht höher ausfallen.

levante Treffer liefert, erreichen Altavista und Fireball "nur" gut ein Drittel, Lycos liegt dazwischen. Im Vergleich der Anzahl relevanter Dokumente schneidet Google.de somit am besten ab.

6.2 Top20 Mean Average Precision

Der Vergleich Top20 (Mikro) Precision mittels Recall-Precision Graph in Abbildung 2 sagt aus, welche Suchmaschine in der Lage ist, die größte Anzahl relevanter Treffer auf den kumulierten Rangplätzen (Cut-off Werte 1-20) über alle Suchanfragen zurückzugeben.

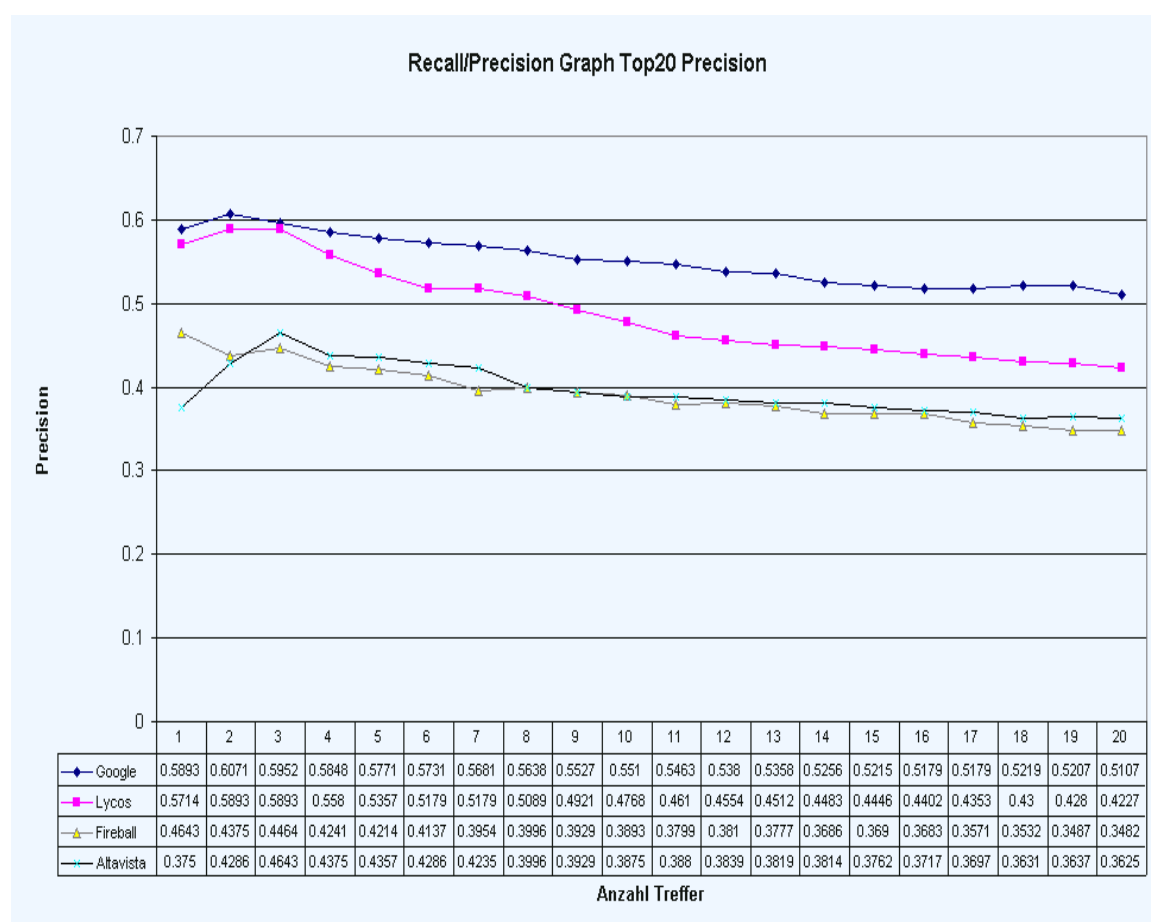


Abbildung 2: Top20 (Mikro-) Precision für google.de, lycos.de, fireball.de und altavista.de

Abbildung 2 ist wie folgt zu deuten. Die Werte auf Rangplatz eins sagen z.B. aus, dass Google über alle Suchanfragen hinweg auf der ersten Position in 58,93% der Fälle relevante Treffer liefert, Lycos 57,14% usw. Die Werte auf Rangplatz zwei sind der Anteil der relevanten Dokumente bis einschließlich Rangplatz zwei. Dies sind z.B. bei Altavista auf Rangplatz zwei 42,86%. Der

Graph kumuliert also immer die Anzahl der relevanten Dokumente bis einschließlich der jeweiligen Position.

Google erreicht die höchste Effektivität, die Top1 Precision beträgt 58,93% und fällt bis auf Platz 20 auf die oben schon genannte Gesamt-, also die Top20 Precision von 51%. Es lässt sich die Aussage treffen, dass Google über alle Suchanfragen bei jedem Cut-off Wert mehr relevante Treffer liefert, als jede andere Maschine. An zweiter Stelle steht Lycos mit einer Top1 Precision von 57% die bis zur Top20 Precision auf 42% fällt. Lycos liefert wiederum zu jedem Cut-off Wert mehr relevante Treffer als Altavista und Fireball. Vergleicht man Altavista und Fireball, so ist nicht unmittelbar klar, welche der beiden Maschinen einen höheren Recall - Precision Graph aufweist. Bei den Rangplätzen 1 und 2 ist Fireball besser, aber ab dem Cut-off Wert von 3 erreicht Altavista mit Ausnahme der Positionen 9 und 10 leicht bessere Werte. Nimmt man als Mean Average Precision²¹ den Durchschnitt der pro Rangplatz erreichten Precisionwerte, um die Resultate der Top1 bis Top20 Precision der verschiedenen Maschinen in einem aussagekräftigen Datum miteinander zu vergleichen, so ergibt sich das Bild in Tabelle 4.

Suchmaschine	Mean Average Precision
Google	0.551
Lycos	0.488
Fireball	0.391
Altavista	0.396

Tabelle 4: Mean Average Precision der Suchmaschinen google.de, lycos.de, fireball.de und altavista.de

Die Überprüfung ergibt, dass Google signifikant besser ist als die anderen Maschinen und Lycos signifikant besser als Altavista und Fireball. Im Vergleich Altavista Fireball lässt sich aber keine klare Aussage treffen.²²

²¹ <http://www-nlpir.nist.gov/works/presentations/spie99/tsld016.htm> [12.04.02].

²² Der Vergleich der kumulierten Precision mittels Vorzeichentest zur Überprüfung der Signifikanz ist in diesem Fall methodisch problematisch, da die Werte der paarweisen Einzelvergleiche nicht unabhängig voneinander sind (z.B. ist der Wert von Top2 zu 50% durch den Wert von Top1 bestimmt). Eine simple alternative Lösung ist aber nicht bekannt, da die Position der Treffer primär durch die Suchmaschinen determiniert wird (Ranking - Treffer auf Platz zwei ist deshalb auf Platz 2, weil er laut Maschine schlechter als der auf Platz 1 und besser als der auf Platz 3 ist). Ein denkbarer anderer Vergleich, etwa nur der Treffer auf den jeweiligen Positionen (z.B. Vergleich der Treffer auf Platz 20 von AV und Google) wäre deshalb grundlegend falsch. Es ist deshalb fraglich, ob die Standardmethodik zur Signifikanzbestimmung abhängiger Messgrößen, Zeitrei-

6.3 Beantwortung der Suchanfragen

Die Retrievaleffektivität der Suchmaschinen bei den einzelnen Suchanfragen (Makroprecision) beschreibt, wie effektiv die Suchmaschinen Informationsbedürfnisse befriedigen.

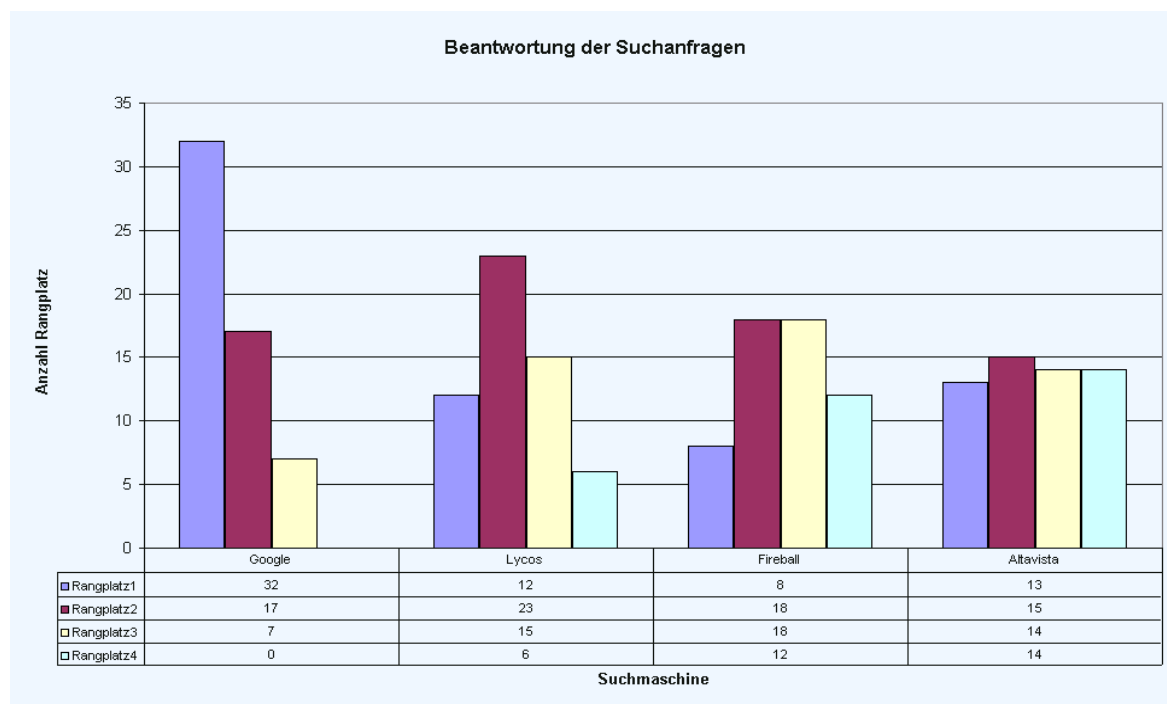


Abbildung 3: Rangplätze in Bezug auf die 56 Suchanfragen der Suchmaschinen google.de, lycos.de, fireball.de und altavista.de

Abbildung 3 zeigt, wie häufig welche Maschine im Vergleich welchen Rangplatz, bezogen auf die Precision bei der jeweiligen Suchanfrage, erreicht. Rangplatz 1 bedeutet, die Maschine erreicht bei der jeweiligen Suchanfrage den höchsten Precisionwert von allen vier Maschinen, Rangplatz 2 den zweithöchsten usw.. Beispiel: bei der Suchanfrage "arbeitsrecht kündigung" erreichte Google eine Precision von 70%, Lycos von 65%, AltaVista von 15% und Fireball eine Precision von 35%. Google erzielt somit Rangplatz 1, Lycos Rangplatz 2, Altavista Rangplatz 4 und Fireball Rangplatz 3.

Google ist auch bei dieser Sichtweise die Maschine mit der höchsten Effektivität. Die Suchmaschine erreicht bei fast der Hälfte der Suchanfragen den höchsten Precisionwert (Rangplatz 1) und belegt bei keiner Suchanfrage Rangplatz 4. Lycos erreicht die zweite Position, da sie häufiger Rangplatz 1 und 2 erzielt als Fireball und Altavista. Die Interpretation ist aber nicht ganz

henanalysen methodisch angemessen oder nicht noch "falscher" wäre. Zumindest sie nur für eine höhere Anzahl von Vergleichsfällen geeignet ist.

eindeutig, denn Altavista erreicht z.B. häufiger Rangplatz 1 als Lycos. Ein intuitiver Vergleich zwischen Fireball und Altavista scheint aufgrund der Ergebnisgrafik nicht sinnvoll, da Altavista alle vier Rangplätze ungefähr gleichhäufig belegt und bei Fireball "das Mittelfeld", also Rangplatz 2 und Rangplatz 3, dominiert.

Die statistische Überprüfung zeigt auch hier, dass Google signifikant besser ist als die anderen Maschinen. Die Unterschiede zwischen Lycos, Altavista und Fireball halten der statistischen Überprüfung hingegen nicht stand.

Insgesamt betrachtet, ist dieses durch den Vergleich der Rangplätze ermittelte Ergebnis eher skeptisch zu bewerten, wie das folgende Beispiel deutlich macht. Bei der Suchanfrage 3 "eheschließungen entwicklung weltweit" erreicht Google eine Precision von 0%, Lycos 10%, Fireball 10% und Altavista 20%. Die vergebenen Rangplätze sind: Google Rangplatz 3, Lycos Rangplatz 2, Fireball Rangplatz 2 und Altavista Rangplatz 1.

Das bedeutet zunächst, dass die Ergebnisse numerisch leicht positiv verfälscht werden, da Rangplätze bei gleich hoher Precision mehrfach vergeben und somit vordere Rangplätze häufiger belegt werden als hintere. Suchanfrage 3 zeigt aber nicht nur diese eher "kosmetische Verzerrung", sondern insbesondere auch ein inhaltlich bedeutendes, weil qualitativ schwerwiegendes Problem auf. Google erzielt bei Suchanfrage 3 eine Precision von 0 und erreicht damit Rangplatz 3. Dabei weist Google zu Suchanfrage 3 keinen relevanten Treffer nach, kann die Suchanfrage also überhaupt nicht beantworten, während die anderen Maschinen zumindest 2 relevante Treffer referenzieren, das Informationsbedürfnis also zumindest graduell befriedigen.

Diese rein quantitative Ergebnisbetrachtung reflektiert also in keiner Weise den grundlegenden qualitativen Gegensatz zwischen „liefert die wenigsten Treffer“ und „kann die Suchanfrage nicht beantworten“. Dabei ist es ein erheblicher Unterschied, ob eine Suchmaschine eine Suchanfrage, wenn auch eher schlecht, beantworten kann oder nicht.

Wie die Problematik dieser sogenannten Nullantworten²³ aufzulösen ist, bleibt ungeklärt. Da dieser Unterschied nicht adäquat quantifiziert werden kann, wird hier darauf verzichtet, diesen, etwa durch die Vergabe von "Strafpunkten", in die Berechnung mit einzubeziehen.

²³ Zur Problematik der Nullantworten siehe [Womser-Hacker 1989], S.151-156, vgl. auch http://www.uni-hildesheim.de/~einf_iw/S10_Eval_IRS.pdf [16.04.02].

Stellt man die Frage, wie viele Suchanfragen die Suchdienste mit zumindest einem relevanten Treffer beantworten können, ergibt sich das Ergebnis aus Tabelle 5. Altavista ist die Maschine, welche die höchste Zahl an Suchanfragen mit zumindest einem relevanten Treffer beantwortet. Google folgt auf Rang zwei und vermag zwei Anfragen weniger zu beantworten.

Suchmaschine	Anzahl beantworteter Suchanfragen (max. 56)	Nicht beantwortet und mehr als 20 Treffer
Google	53	2
Lycos	50	2
Fireball	51	2
Altavista	55	1

Tabelle 5: Anzahl beantworteter Suchfragen bzw. ohne relevante Treffer

7 Bewertung der Ergebnisse

In einem Satz zusammengefasst lautet das Ergebnis dieser Untersuchung: Google ist effektiver als die anderen Suchmaschinen. Zwischen den anderen Maschinen lässt sich kein Unterschied bei der Retrievaleffektivität feststellen. Zu beachten bleibt, dass dieses Ergebnis sich nur auf die zugrundeliegenden Zahlenwerte stützt, d.h. davon ausgeht, dass diese objektiv ermittelt werden konnten. Die Verzerrungen, die sich durch die zeitliche Divergenz zwischen Anfrage- und Beurteilungszeitpunkt sowie die Nulltreffermengen ergeben, bleiben unberücksichtigt.

Die Problematik der Nulltreffersuchanfragen, zeigt sehr deutlich, dass die je nach Bewertungssicht und Bewertungsmaß variierenden Effektivitätswerte immer von den Faktoren präjustiert werden, die dem jeweiligen Retrievaltest immanent sind.²⁴

Wie sind die Ergebnisse dieses Testes also einzuordnen? Den normativen Rahmen dieser Untersuchung bilden primär die Thematik, Komplexität und Spezifität der Fragestellungen. Suchanfragen und Informationsbedürfnisse, die freizeitbezogen sind, Produkt- und Dienstleistungsangebote kommerzieller Anbieter zum Ziel haben oder erotischer/pornographischer Natur sind, werden ausgeschlossen. Die Suchanfragen dieser Untersuchung sind thematisch eher dem (sozial)wissenschaftlichen, politischen Umfeld zuzuordnen. Die Ergebnisse dieser Untersuchung gelten nur für solche Informationsbedürfnisse und Suchanfragen.

²⁴ Untersuchungsziel, Bewertungsmaße, Kriterien der Relevanzbeurteilung, Art und Ausgestaltung der Informationsbedürfnisse, Durchführung der Suchanfragen.

Diese Ergebnisse beruhen dabei ausschließlich auf der Beurteilung der Relevanz der ersten 20 Ergebnisseiten. Mehrwerte, wie die Hilfe zur Vorab-Relevanzbeurteilung durch die Metainformation der Trefferliste der Suchmaschinen, das Browsing in thematisch passenden Rubriken²⁵ oder der Zugriff auf nicht mehr vorhandene Webseiten mit Hilfe des Google Archivs bleiben in der Testanordnung unberücksichtigt. Die Suchmaschine Google bietet dabei mehrere solcher Mehrwerte als Alleinstellungsmerkmale. Beispiele hierfür sind u.a. Zugriff auf PDF- und Postscript-Dokumente oder der Nachweis nicht mehr verfügbarer Dokumente über das Google Archiv. Da diese Alleinstellungsmerkmale den Rechercheerfolg bei Google im Vergleich zu anderen Maschinen zumindest tendenziell begünstigen, ist davon auszugehen, dass Google für den Nutzer in weitaus größerem Maße die beste Alternative unter den untersuchten Suchmaschinen darstellt, als die Ergebniswerte allein vermuten lassen.

8 Einschätzung des Retrievaltests

Die Zielsetzung dieses Retrievaltests ist es, die Retrievaleffektivität der untersuchten Suchmaschinen zu ermitteln und das in [Griesbaum 2000] konzipierte Evaluationsverfahren zu optimieren und weiterzuentwickeln. Zunächst kann festgehalten werden, dass die auf [Tague-Sutcliffe 1992] beruhende methodische Konzeption geeignet erscheint, den Retrievaltest strukturiert zu konzipieren und bei auftauchenden Schwierigkeiten Problemlösungsstrategien vor der Testdurchführung zu entwickeln. Die von [Hawking et al. 2001] genannten Rahmenbedingungen für die Evaluation von Suchmaschinen können eingehalten werden. Es werden reale Informationsbedürfnisse eingebunden, die durch eine große Anzahl von Suchfragen repräsentiert wurden. Ebenso werden die wichtigsten Suchmaschinen für den deutschen Sprachraum in der Untersuchung berücksichtigt. Inwieweit Suchfrageformulierungen ein Informationsbedürfnis gut repräsentieren, bleibt natürlich immer schwer zu beantworten. Aber durch die meist kurze Frageformulierung ist zumindest die Länge der Suchfragen den Gegebenheiten im Web angepasst.

Die Optimierung des Evaluationsverfahrens soll in dieser Untersuchung durch die Entwicklung und Verwendung eines automatischen Abfrageskripts umgesetzt werden, das im Vergleich zur „händischen“ Durchführung der Suchanfragen in der genannten Untersuchung, durch die Minimierung der zeitlichen Verzerrungen zwischen Anfrage- und Bewertungszeitpunkt, zugleich Validi-

²⁵ Beispielsweise bei den Katalogtreffern von Lycos und Google.

täts- und Effizienzsteigerungen bewirken soll. Dieses Ziel wird nicht vollständig erreicht, da der Aufwand, ein solches Skript zuverlässig auszugestalten und persistent am Laufen zu halten, sehr hoch ist. Da die Suchmaschinen zu jedem Zeitpunkt Modifikationen durchführen, z.B. die Abfrage-URL ändern können, kann letztlich keine Garantie dafür gegeben werden, dass es zum benötigten Zeitpunkt überhaupt funktioniert. Aus diesem Grund werden bei der Testdurchführung aus Sicherheitserwägungen die Suchanfragen im Voraus durchgeführt und somit im Maximalfall zwei Tage zeitlicher Verzögerung zwischen Suchanfragendurchführung und Relevanzbeurteilung bewusst in Kauf genommen.

Die Nulltrefferproblematik zeigt, dass die Validität der Untersuchungsergebnisse auch von den Ergebniswerten selbst abhängig ist. Hätten beispielsweise alle Suchmaschinen zu allen Anfragen mindestens einen relevanten Treffer geliefert, so wäre dieses Problem gar nicht aufgetreten. Prinzipiell ist zu fragen, inwieweit das Ziel, die Testparameter realitätsnah auszugestalten, durch die Annäherung an sogenannte typische Verhaltensmuster von Suchdienstnutzern erreicht werden kann, wenn diese typischen Verhaltensmuster durch die quantitative Generalisierung qualitativ unspezifizierter Interaktionsdaten, z.B. durch Logfileanalyse, identifiziert werden und gerade deshalb nicht zutreffen müssen. Hier stellt sich wieder das für Retrievaltests zentrale Problem der qualitativen Ausgestaltung der Testparameter, für das keine theoretische Lösung vorliegt. Eine einfache Lösung ist nicht abzusehen, umso dringender ist der Bedarf, Verfahren und Methoden zu entwickeln, die solche qualitativen Probleme bei scheinbar objektivierbaren Kriterien aufdecken.

Hinzu kommt ein weiteres Problem, welches die Annahme, dass die Herkunft der Treffer mit vertretbarem Aufwand verschleiert werden kann, grundsätzlich in Frage stellt. Die Alleinstellungsmerkmale von Google.de hinsichtlich indexierter Dokumentformate (PDF, Postscript) werden zwar in die Testanordnung mit integriert, aber beim Google Archiv gelingt das nicht. Hier stellt sich also die Frage, ob die grundlegende Testanforderung, die Herkunft der Ergebnisse unkenntlich zu machen, weiterhin durchzuhalten ist oder durch Weiterentwicklungen der Suchmaschinen bei der Ergebnispräsentation künftig nicht obsolet wird.²⁶

Fasst man diese Probleme und Schwierigkeiten zusammen, lässt sich Folgendes festhalten:

²⁶ Vgl. die Visualisierung der Suchergebnisse bei dem Metadiens KartOO.
URL <http://www.kartoo.com> [11.05.02].

Das bei diesem Retrievaltest verfolgte Ziel, die Validität und Effizienz der Testanordnung durch die Entwicklung eines automatischen Anfrageskript zu erhöhen, wird aus technischen Gründen nicht erreicht, die Effizienz jedoch erhöht. Im Laufe der Untersuchung zeigt sich, dass sich die realitätsnahe Ausgestaltung der Testparameter weit schwieriger gestaltet, als in der Vorgängeruntersuchung und weitaus weniger umgesetzt werden kann. Dies liegt zum einen in Ergebniswerten begründet, die qualitative Annahmen des Testdesigns in Zweifel ziehen (Top 20 Grenze), hat zum anderen aber auch ihre Ursache in den Merkmalen von Suchmaschinen (Google Cache), die Grundanforderungen der Testanordnung (Unkenntlichmachung der Treffer) grundlegend in Frage stellen.

Aus den Problembereichen der Untersuchung wird ein breites Optimierungspotenzial bei künftigen Untersuchungen, insbesondere bezüglich der Durchführung der Suchanfragen, der Anzahl der bewerteten Dokumente und der Ergebnisdarstellung ersichtlich. Daraus lässt sich schlussfolgern, dass bei künftigen Untersuchungen mehr Ressourcen für die realitätsnahe Ausgestaltung der Testparameter bereitgestellt werden sollen.²⁷

9 Literatur

- Bager, J. & Schulzki-Haddouti, C. (2001). Alle gegen Google. Wisenut, Teoma, Vivisimo – neue Konkurrenz für die populäre Suchmaschine. *c't Magazin für Computer Technik* Nr.19, 104-108
- Bar-Ilan, J. (2002). Methods for measuring search engine performance over time. *Journal of the American Society for Information Science and Technology*, 53 Nr.4, 308-319
- Buckley, C. & Voorhees, E. M. (2000). Evaluating Evaluation Measure Stability. In: *SIGIR 2000*. Belkin, N. J., Ingwersen, P., und Leong, M.-K. (eds.); ACM, 33-40
- Dennis, S., Bruza, P., und McArthur, R. (2002). Web searching: a process-oriented experimental study of three interactive search paradigms. *Journal of the American Society for Information Science and Technology*, 53 Nr.2, 120-133
- Dresel, R., Hörnig, D., Kaluza, H., Peter, A., Roßmann, N., und Sieber, W. (2001). Evaluation deutscher Web-Suchwerkzeuge. Ein vergleichender Retrievaltest. *nfd Information – Wissenschaft und Praxis*, 52 Nr.6, 381-392
- Ellis, D. (1992). Paradigms and proto-paradigms in information retrieval research. In: *Conceptions of library and information science: historical, empirical, and theoretical perspectives*. Vakkari, P. and Cronin, B. (eds.); London: Taylor Graham, 165-186

²⁷ Die Autoren bedanken sich bei den Studierenden der Informationswissenschaft in Konstanz und Düsseldorf, die an dem Retrievaltest mit viel Interesse und Engagement teilgenommen haben. Besonderer Dank gebührt Caroline Berns, Marion Herb und Birgit Scherer, die im Rahmen studentischer Projekte an der Durchführung und Auswertung beteiligt waren.

- Ford, N., Miller, D., und Moss, N. (2001). The role of individual differences in Internet searching: an empirical study. *Journal of the American Society for Information Science*, 52 Nr.12, 1049-1066
- Gordon, M. & Pathak, P. (1999). Finding information on the World Wide Web: the retrieval effectiveness of search engines. *Information Processing & Management*, 35 Nr.2, 141-180
- Griesbaum, J. (2000). Evaluierung hybrider Suchsysteme im WWW.
http://www.inf.uni-konstanz.de/~griesbau/files/evaluierung_hybrider_suchsysteme_im_www.pdf
22.04.2002
- Hawking, D., Craswell, N., Bailey, P., und Griffiths, K. (2001). Measuring search engine quality. *Information Retrieval*, 4 Nr.1, 33-59
- Jansen, B. J., Spink, A., und Saracevic, T. (2000). Real Life, Real Users, and Real Needs: A Study and Analysis of User Queries on the Web. *Information Processing & Management*, 36 Nr.2, 207-227
- Körber, S. (2000). Suchmuster erfahrener und unerfahrener Suchmaschinennutzer im deutschsprachigen World Wide Web. Ein Experiment.
<http://kommunix.uni-muenster.de/IfK/examen/koerber/>
- Leighton, H. V. & Srivastava, J. (1999). First 20 precision among World Wide Web search services (search engines). *Journal of the American Society for Information Science*, 50 Nr.10, 870-881
- Lesk, M. (1995). The seven ages of information retrieval. In: Conference for the 50th anniversary of As We May Think. MIT, 12-14
- Mettrop, W. (2001). Internet search engines-fluctuations in document accessibility. *Journal of Documentation*, 57 Nr.5, 623-651
- o.V. (2001). Internet-Suchmaschinen. Google trifft am besten. Test.Stiftung Warentest Nr.9, <http://www.warentest.de>
- Oppenheim, C., Morris, A., McKnight, C., und Lowley, S. (2000). Progress in documentation the evaluation of WWW search engines. *Journal of Documentation*, 56 Nr.2, 190-211
- Robertson, S. E. (1981). The methodology of information retrieval experiments. In: *Information Retrieval Experiment*. Jones, K. Sparck (ed.); London: Butterworth, 9-31
- Stock, M & Stock, W. (2000). Internet-Suchwerkzeuge im Vergleich. Retrievaltest mit Known Item Searches. *Password*, 11, 23-31
- Su, L. T. & Chen, H.-l. (1999). Evaluation of Web Search Engines by Undergraduate Students. In: *Knowledge: Creation, Organization and Use. ASIS '99 – Proceedings of the 62nd ASIS Annual Meeting*. Woods, L. (ed.); Information Today, Inc., 98-114
- Tague-Sutcliffe, J. (1992). The pragmatics of information retrieval experimentation, revisited. *Information Processing & Management*, 28 Nr.4, 467-490
- Voorhees, E. M. & Harman, D. (2001). Overview of TREC 2001. 2001,
http://trec.nist.gov/pubs/trec10/papers/overview_10.pdf, 01.05.02
- Wang, H., Xie, M., und Goh, T. N. (1999). Service quality of Internet search engines. *Journal of Information Science*, 25 Nr.6, 499-507

- Warner, J. (2000). In the catalogue ye go for men: evaluation criteria for information retrieval systems. ASLIB Proceedings, 52 Nr.2, 76-82
- Wolff, C. (2000). Effektivität von Recherchen im WWW. Vergleichende Evaluierung von Such- und Metasuchmaschinen. In: Informationskompetenz – Basiskompetenz in der Informationsgesellschaft. Proceedings des 7. Internationalen Symposiums für Informationswissenschaft. Knorz, G. & Kuhlen, R. (eds.); Universitätsverlag Konstanz, 31-48
- Womser-Hacker, C. (1989). Der PADOK Retrievaltest. Zur Methode und Verwendung statistischer Verfahren bei der Bewertung von Information-Retrieval-Systemen. Hildesheim: Georg Olms



Das Cross-Language Evaluation Forum (CLEF) - Evaluationsumgebung und Forschungskontext für mehrspra- chiges Information Retrieval (mit einer Skizze der Ergebnisse von CLEF 2002)

Michael Kluck

Informationszentrum Sozialwissenschaften (IZ), Bonn /
Humboldt-Universität zu Berlin, Abteilung Pädagogik und Informatik
michael.kluck@educat.hu-berlin.de

Zusammenfassung

Einleitend werden die Struktur und Methodik von CLEF kurz erläutert. Über ausgewählte Ergebnisse der CLEF-Kampagne 2002 wird berichtet, insbesondere über die verwendeten Verfahren und Ressourcen für die Behandlung der Multilingualität sowie über das interaktive und das fachbezogene Cross-Language Information Retrieval (CLIR). Die Möglichkeiten zur Beteiligung an der Forschung im Bereich des mehrsprachigen Information Retrieval (CLIR) und der Methodologie der IR-Evaluationen werden aufgezeigt. Abschließend wird die weitere Entwicklung von CLEF skizziert.

Abstract

At the beginning structure and methodology of CLEF are briefly explained. Selected results of the CLEF 2002 campaign will be reported, especially the used procedures and resources for the treatment of multilinguality and the interactive and scientific cross-language information retrieval (CLIR). The possible fields of research in the area of CLIR and its evaluation methodology are shown. Finally the further development of CLEF is shortly described.

1 Einleitung: CLEF im Kontext der IR-Forschung

Das Cross-Language Evaluation Forum (CLEF) ist als ein Spin-off der TREC-Evaluierungskampagne aus dem damaligen TREC-Task Cross-Language Information Retrieval (CLIR) entstanden (näheres siehe Harman et al. 2001 und Kluck et al. 2002), das sich schwerpunktmäßig mit dem CLIR europäischer Sprachen befasst. Parallel dazu hat sich auch eine entsprechende



Initiative (NTCIR)¹ in Japan entwickelt, die sich auf die fernöstlichen Sprachen (Japanisch, Chinesisch, Koreanisch etc.) konzentriert, während sich TREC nun mit den arabischen Sprachen befasst. Alle drei Initiativen fußen auf der bei TREC weiter entwickelten Cranfield-Methodologie.

Die Anforderungen an das mehrsprachige Information Retrieval (CLIR) gehen über das des einsprachigen IR insofern hinaus, als die Transponierung der Anfrage aus einer Sprache in die weiteren Sprachen der durchsuchten Dokumentmengen erforderlich ist und eine integrierte gerankte Antwortliste mit verschiedensprachigen Ergebnisdokumenten gefordert wird.

2 CLEF – Struktur und Methoden der Evaluation

2.1 Verwendete Datenbestände und Sprachen der Dokumente

Die Datenbestände, die in CLEF Verwendung finden, umfassen vorwiegend Zeitungsartikel und Meldungen von Nachrichtenagenturen. Jeder Teilbestand in den verschiedenen Sprachen liegt für das Jahr 1994 vollständig vor. Dadurch soll sichergestellt werden, dass die gleichen Themenkomplexe bzw. Ereignisse in den verschiedenen Sprachen behandelt werden. Die Kernsprachen sind Deutsch Englisch, Französisch, Italienisch und Spanisch. Ferner stehen Datenbestände in Finnisch, Holländisch und Schwedisch zur Verfügung. Für das fachbezogene CLIR stehen Datenbestände aus Fachdatenbanken in Französisch und Deutsch zur Verfügung.

Die folgenden Korpora stehen zur Verfügung:

- Zeitungen und Nachrichtenagenturen in den Kernsprachen
- Englisch (Los Angeles Times 1994): 113 005 Dokumente, 425 MB
- Deutsch (Schweizerische Depeschenagentur – SDA 1994, Der Spiegel 1994-1995, Frankfurter Rundschau 1994): 225 371 Dokumente, 527 MB
- Französisch (Schweizerische Depeschenagentur – SDA 1994, Le Monde 1994): 87 191 Dokumente, 243 MB
- Italienisch (Schweizerische Depeschenagentur – SDA 1994, La Stampa 1994): 108 578 Dokumente, 278 MB
- Spanisch (Agencia EFE 1994): 215 738 Dokumente, 509 MB

¹ NTCIR (NII-NACSIS Test Collection for IR Systems); NII = National Institute of Informatics, Tokyo, Japan; NACSIS = National Center for Science Information Systems

wissenschaftliche und domänenspezifische Daten

- Wissenschaften (alle Fachgebiete): AMARYLLIS (Französisch): 148 688 Dokumente, 195 MB
- Sozialwissenschaften: GIRT (Deutsch): 76 128 Dokumente, 150 MB

weitere Daten für bilinguale Tests (Zeitungen und Nachrichtenagenturen)

- Holländisch (NRC Handelsblad 1994-1995, Algemeen Dagblad 1994-1995): 190 604 Dokumente, 540 MB
- Finnisch (Aamulehti Ende-1994-1995) 55 344 Dokumente, 137 MB
- Schwedisch: (Tidningarnas Telegrambyrå 1994-1995) 142 819 Dokumente, 352 MB

Diese Dokumentsammlungen umfassen in der Regel das Jahr 1994 und teilweise auch 1995. Die einzelnen Dokumente der Korpora sind mit SGML-Tags für die Auszeichnung der enthaltenen Datenelemente versehen.

2.2 Themenstellungen (Topics)

Die Erarbeitung der Themenstellungen (Topics) für CLEF ist ein komplexer Prozess, an dem die verschiedenen CLEF-Sprachgruppen (DE, EN, ES, FR, IT, NL, FI, SV) intensiv beteiligt sind. Zusätzlich werden die Themenstellungen von weiteren Gruppen in andere (auch außereuropäische) Sprachen übersetzt (ZH, JP etc.). Die Sprachgruppen müssen Themenstellungen entwickeln, die dem Inhalt der Dokumente entsprechen, die von den verschiedenen Zeitungen und Nachrichtenagenturen im Jahr 1994/1995 gemeldet wurden: z.B. Ereignisse aus Politik, Kultur, Sport, Wissenschaft, Medizin. Jede Sprachgruppe schlägt bis zu 15 Themen vor. Aus diesen mehr als 80 Themenvorschlägen werden schließlich 50 Themenstellungen ausgewählt, die dann als Topics den Teilnehmern der Evaluationskampagne zur Verfügung gestellt werden. Dann werden die 50 Topics aus der jeweiligen Originalsprache in die übrigen Sprachen übersetzt. Schließlich wird von einer unabhängigen Gruppe von Fachübersetzern eine Nachprüfung der Übersetzungen in alle offiziellen Sprachen vorgenommen (topic check), um die Konsistenz und Genauigkeit der Übersetzungen zu sichern.

Für die wissenschaftlichen bzw. domänenspezifischen Aufgabenstellungen im Rahmen von GIRT² (German Indexing and Retrieval Database) und AMARYLLIS werden jeweils 25 fachspezifische Themen auf Deutsch bzw. Französisch entwickelt. Für diese wird auch eine englische (und bei GIRT eine

² GIRT stellt ein Korpus sozialwissenschaftlicher Fachtexte zur Verfügung.

russische Übersetzung) vorbereitet, so dass auch hier CLIR Ansätze getestet werden können.

2.3 Topic Beispiele jeweils mit einer Übersetzung

Die Topics sind mit SGML-Tag versehen und bestehen aus einer fortlaufenden Nummer, dem Titel (zwei bis drei Worte zur Kurzcharakterisierung des Themas), der Beschreibung (ein ganzer Satz zur Bezeichnung des Themas) und der ausführlichen Beschreibung (mit Erläuterungen, welche Inhalte einschlägig sind und welche möglicherweise ausgeschlossen werden sollen).

Beispiel 1: Rinderwahnsinn

```
<top>
<num> C088 </num>
<EN-title> Mad Cow in Europe </EN-title>
<EN-desc> Find documents that cite cases of Bovine Spongiform
Encephalopathy (the mad cow disease) in Europe. </EN-desc>
<EN-narr> Relevant documents will report statistics and/or figures on cases
of animals infected with Bovine Spongiform Encephalopathy (BSE),
commonly known as the mad cow disease, in Europe. Documents that only
discuss the possible transmission of the disease to humans are not considered
relevant. </EN-narr>
</top>
<top>
<num> C088 </num>
<IT-title> Mucca pazza in Europa </IT-title>
<IT-desc> Trova i documenti che citano i casi di mucca pazza (Encefalopatia
Spongiforme Bovina) in Europa. </IT-desc>
<IT-narr> Sono rilevanti i documenti che riportano statistiche e/o dati
numerici sui casi di animali affetti da Encefalopatia Spongiforme Bovina
(BSE), comunemente detta morbo della mucca pazza, in tutti i paesi europei.
Non sono rilevanti i documenti sulla possibile trasmissione del morbo
all'uomo. </IT-narr>
</top>
```

Beispiel 2: Mausarm

<top>
<num> C064 </num>
<NL-title> Muisarm </NL-title>
<NL-desc> Zoek documenten waarin melding wordt gemaakt van een muisarm. </NL-desc>
<NL-narr> Relevante documenten melden klachten die veroorzaakt worden door het langdurig gebruik van een computermuis. Documenten waarin maatregelen worden genoemd voor het voorkomen van een muisarm tijdens het werken achter de computer zijn ook relevant. </NL-narr>
</top>
<top>
<num> C064 </num>
<ES-title> Síndrome RSI y ratones de ordenador </ES-title>
<ES-desc> Encontrar documentos que informen sobre RSI ("repetitive strain injuries" o "enfermedad del periodista") producidas por el uso del ratón del ordenador. </ES-desc>
<ES-narr> Los documentos relevantes informan sobre daños causados por el uso continuado de un ratón de ordenador. Los documentos que proponen formas de evitar el RSI cuando se usa el ordenador también son relevantes. </ES-narr>
</top>

Beispiel 3: Eurofighter

<top>
<num> C093 </num>
<DE-title> Eurofighter </DE-title>
<DE-desc> Finde Dokumente, die über das Projekt "Jäger 90" bzw. den "Eurofighter" berichten. </DE-desc>
<DE-narr> Relevante Dokumente berichten über das Projekt "Jäger 90". Die Partnerländer Deutschland, Großbritannien, Italien und Spanien hatten das "Eurofighter"-Konsortium gebildet. Von Interesse sind auch der Arbeitsanteil der beteiligten Rüstungskonzerne sowie Kostenschätzungen. </DE-narr>
</top>

<top>
<num> C093 </num>
<ES-title> Eurofighter </ES-title>

<ES-desc> Encontrar documentos que informen acerca del proyecto EFA o "Eurofighter". </ES-desc>
<ES-narr> Los documentos relevantes informan sobre el proyecto EFA. Los países socios, Alemania, Reino Unido, Italia y España formaron el consorcio "Eurofighter". Son también de interés las informaciones sobre el reparto de trabajo entre los grupos involucrados, así como las estimaciones de costes. </ES-narr>
</top>

Beispiel 4: Erster finnischer EU-Kommissar

<top>
<num> C118 </num>
<FR-title> Le premier commissaire européen finlandais à l'UE </FR-title>
<FR-desc> Quel personnage d'origine finlandaise a été élu commissaire européen à l'Union Européenne? </FR-desc>
<FR-narr> Donner le nom du premier Finlandais nommé commissaire européen. Les documents pertinents peuvent également faire mention des tâches et missions qui incomberont au nouveau commissaire. </FR-narr>
</top>

<top>
<num> C118 </num>
<FI-title> Suomen ensimmäinen EU-komissaari </FI-title>
<FI-desc> Kuka nimettiin Suomen ensimmäiseksi komissaariksi Euroopan Unioniin? </FI-desc>
<FI-narr> Kerro Suomen ensimmäisen Euroopan Unionin komissaarin nimi. Relevantteja ovat myös dokumentit jotka kertovat uuden komissaarin toimeen liittyvistä toimintaaalueista. </FI-narr>
</top>

2.4 Pooling-Methode

Die Schritte des Bewertungsverfahrens, das auf der Pooling-Methode von TREC basiert, sind wie folgt: Die teilnehmenden Systeme liefern ihre integrierten, geordneten Ergebnislisten pro Themenstellung. Diese Listen enthalten die Nummern derjenigen Dokumente für die jeweiligen Themen, die von den Systemen als relevant ermittelt wurden, und zwar in der Reihenfolge der vermuteten Relevanz (absteigend). Die jeweils ersten 60 Dokumente werden in den Pool für den entsprechenden Topic einbezogen. Alle Ergebnislisten, die

zu einer der 50 Themenstellungen der Hauptaufgabe bzw. zu einer der je 25 Themenstellungen der domänenspezifischen Aufgabe (GIRT) und der wissenschaftlichen Aufgabe (MARYLLIS) gehören, werden zusammengespielt und in eine Zufallsreihenfolge³ gebracht; dadurch ist nicht mehr feststellbar, welches Dokument von welchem System stammt oder an welcher Stelle der vermuteten Relevanzreihenfolge es stand. Dann werden die Listen nach Sprachen aufgeteilt: alle Dokumentnummern aus den Korpora, die zu einer bestimmten Sprache gehören, werden zusammengeführt. Auf diese Weise wird pro Themenstellung eine große Zahl von Dokumenten zur jeweiligen Sprache zur Relevanzbewertung gesammelt. Die zu bewertende Maximalanzahl von Dokumenten könnte $n \cdot 60 \cdot 50$ ⁴ sein, aber tatsächlich kommt nur etwas mehr als ein Viertel dieser Anzahl zusammen, weil nicht alle Gruppen die Maximalanzahl von Treffern pro Thema und Sprache beitragen und manche Treffer von mehreren Systemen abgeliefert werden - also identisch sind. Für die beiden wissenschafts- bzw. fachbezogenen Aufgabenstellungen kann maximal jeweils eine separate Anzahl von $n \cdot 60 \cdot 25$ erwartet werden, aber auch hier wird de facto nur etwas mehr als ein Viertel dieser Anzahl erreicht.

2.5 Relevanz-Bewertung

Diese gemeinsamen, geordneten Ergebnislisten pro Sprache werden von den Juroren der jeweiligen Sprachgruppe beurteilt. Die Beurteilung wird mit Hilfe der von NIST entwickelten Bewertungssoftware ASSESS festgehalten. Die Entscheidungen der Bewerter (als relevant oder nicht relevant in Hinsicht auf das jeweilige Thema) werden den sprachbezogenen Ergebnislisten jeder Themenstellung hinzugefügt. Die Bewerter wenden allgemeine Beurteilungsregeln an, die mit denjenigen von TREC vergleichbar sind; sie ziehen die Themendiskussionen der Sprachgruppen als Richtlinien für die Relevanzentscheidungen in Betracht und sie benutzen die Narratives der Themenstellungen als Entscheidungshilfen.

Trotzdem ist die binäre Entscheidung manchmal schwierig und häufig äußern die Juroren den Wunsch nach einer abgestuften Skala von Relevanz. Wie TREC hält aber auch CLEF aufgrund der besseren Auswertbarkeit an der binären Relevanzentscheidung fest.

³ Dies ist die aufsteigende Reihenfolge nach den Dokumentnummern der jeweiligen Korpora.

⁴ n (Anzahl von teilnehmenden Gruppen)
60 (zutreffende relevante Dokumente pro Themenstellung)
50 (Anzahl von Themenstellungen).

2.6 Aufgabenstellungen (Tasks)

CLEF zielt auf die Weiterentwicklung von multilingualen IR-Systemen ab. Auf dem Weg zum Testen mehrsprachiger IR-Systeme sollen jedoch auch Schritte zur Einbeziehung weiterer Sprachen und zum Sammeln von Erfahrungen bei der Ausrichtung von Tests möglich sein. Daher werden unterschiedliche Aufgabestellungen (Tasks) formuliert, denen sich die Teilnehmer stellen können⁵.

Die Hauptaufgabenstellung von CLEF ist das **mehrsprachige IR** (multilingual task):

Suche nach Dokumenten in allen Hauptsprachen, wobei eine dieser Sprachen als Anfragesprache dient, sowie Erstellung einer integrierten Liste aller Ergebnisse aus allen Dokumentsammlungen (d.h. aus allen Hauptsprachen).

Zusätzlich werden weitere Sprachen als Ausgangssprachen möglich gemacht, da entsprechende Übersetzungen der Themenstellungen von teilnehmenden Gruppen erstellt werden (Finnisch, Russisch, Schwedisch). Die Hauptsprachen bleiben dabei ebenfalls Zielsprachen.

Die **zweisprachige Aufgabenstellung** (bilingual task) sieht vor, in einer beliebigen Ausgangssprache (die nicht gleich der Zielsprache ist) nach Dokumenten in englischer oder holländischer Sprache zu suchen. Aus diesem Grund stellten die CLEF-Organisatoren auch Übersetzungen der Themenstellungen ins Holländische sowie weitere linguistische Ressourcen für das Holländische (Stoppwortliste, Stemmer, Holländisch-Englisch Lexikon) zur Verfügung.

Die **einsprachige Aufgabenstellung** (monolingual task) sieht vor, in deutscher, englischer, französischer, holländischer, italienischer und spanischer Sprache nach Dokumenten in einer der entsprechenden Dokumentsammlungen zu suchen. Englisch wird dabei ausgeschlossen, da Englisch schon in der Vergangenheit durch den Ad-hoc-Retrieval Task von TREC abgedeckt wurde und keine neue Herausforderung im Hinblick auf linguistische Probleme und Übersetzungsfragen darstellt. Die einsprachige Aufgabenstellung ist einerseits als Einstieg für neue CLEF-Teilnehmer gedacht, andererseits können auf diese Weise neue Sprachen für die multilinguale Aufgabenstellung eingeführt werden.

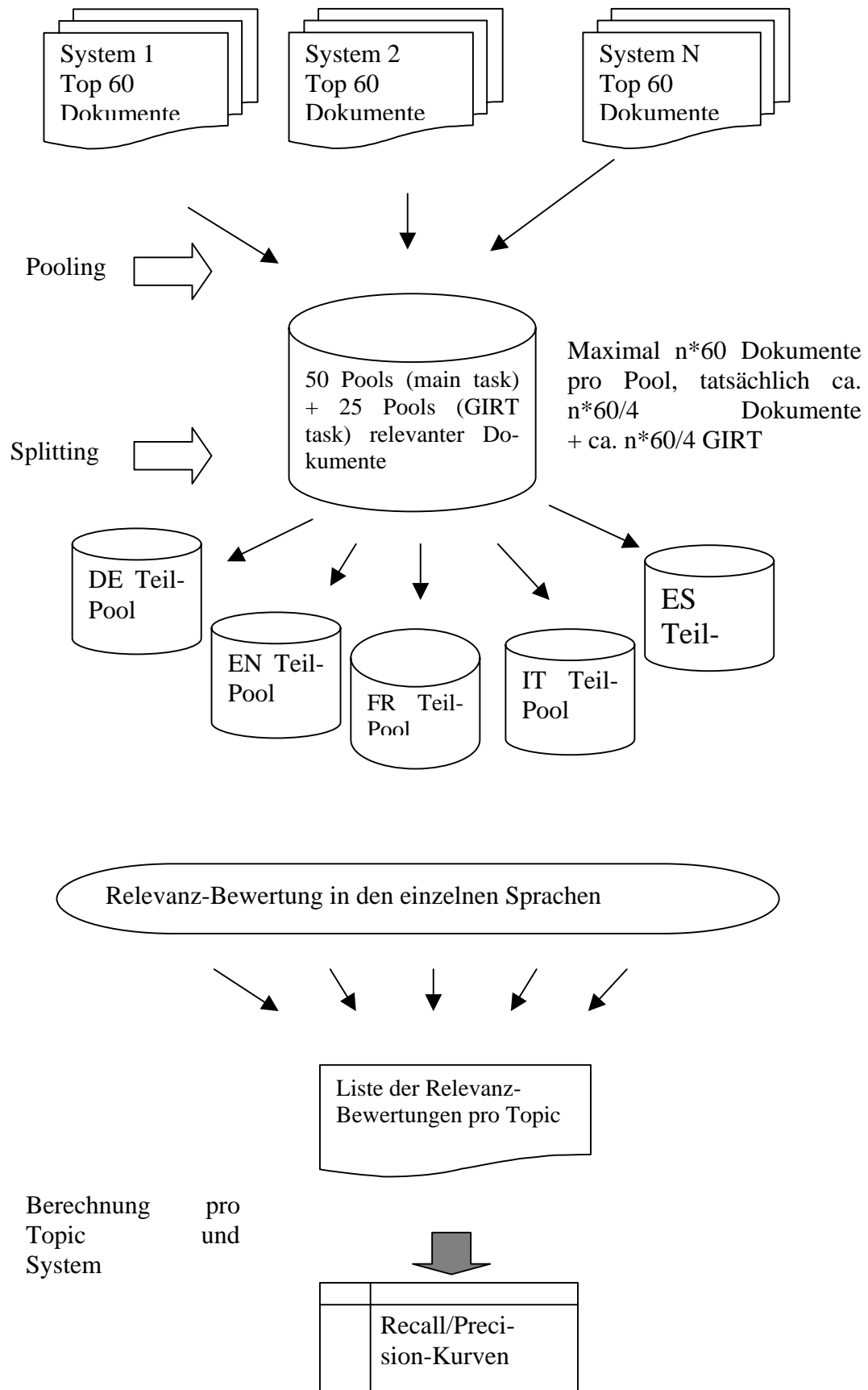
⁵ <http://www.iei.pi.cnr.it/DELOS/CLEF/clef00.html>.

Die wissenschaftliche bzw. **fachbezogene Aufgabenstellung** (scientific and domain-specific task) erlaubt, nach (sozial)wissenschaftlichen Dokumenten in speziellen Dokumentensammlungen, nämlich GIRT (German Indexing and Retrieval Testdatabase) oder AMARYLLIS zu suchen. Damit reagiert CLEF auf den oft genannten Vorwurf, große Evaluierungen nur auf der Basis von Zeitungstexten führten nicht zu übertragbaren Ergebnissen. Die Dokumente der GIRT- und AMARYLLIS -Datenbanken enthalten auch intellektuell vergebene Schlagwörter jeweils aus einem (sozial)wissenschaftlichen Thesaurus, der ebenfalls zur Verfügung gestellt wird, und zwar auch in englischer bzw. bei GIRT zusätzlich noch in russischer Übersetzung. Außerdem werden dafür spezifische Themenstellungen auf Englisch und Deutsch bzw. Französisch (GIRT außerdem Russisch) bereitgestellt. AMARYLLIS und GIRT bieten eine ideale Plattform, um die Übertragbarkeit der Systeme auf spezifische wissenschaftliche Fachtexte zu testen (cf. Gey & Kluck 2001).

Zusätzlich wurde eine experimentelle Aufgabenstellung definiert: die **interaktive Aufgabenstellung** (interactive task). Ziel dieser Aufgabe ist es, die Evaluierung von interaktivem CLIR zu erforschen und Vergleichsmaßstäbe zu entwickeln, an denen weitere Forschungen gemessen werden können. In diesem Track wird also die Retrievaleffektivität in Kombination mit der Benutzungsoberfläche bewertet. Dabei geht es insbesondere um die Möglichkeiten, die Anfrage zu formulieren und zu verändern und die Ergebnisdokumente schnell bewerten zu können. Die Anfragen werden in diesem Fall von Testpersonen bearbeitet und nicht automatisch vom System oder von Experten erstellt (cf. Gonzalo & Oard 2002).

2.7 Bearbeitung der Themenstellungen durch die Teilnehmer

Die von den Teilnehmern verwendeten Retrievalsysteme beginnen mit der Suche nach einer Themenstellung in einer Sprache (z.B. DE) und liefern Dokumente in allen Zielsprachen zurück. Für die Suchen in den Dokumentbeständen (in der Hauptaufgabenstellung des mehrsprachigen IR: DE, EN, ES, FR, IT) verwenden sie systemspezifische Strategien, um die Übersetzung oder Transformation der Fragen in andere Sprachen zu lösen. Am Ende des Rückgewinnungsprozesses müssen sie einen integrierten und geordneten Ergebnissatz der Dokumente liefern, von denen angenommen wird, dass sie für die Themenstellung relevant sind. Der Prozess der Integration der Ergebnisse aus verschiedenen Datenbeständen (merging) ist neben der Lösung der Übersetzungsprobleme eine zweite nicht unwichtige Herausforderung.



3 Die CLEF-Kampagne 2002: Überblick

Der Überblick wird die folgenden Abschnitte behandeln und aktuell vom CLEF 2002 Workshop im September 2002 in Rom berichten:

- Teilnehmer und verwendete Sprachen
- verwendete Verfahren zum CLIR
- verwendete Ressourcen für die Bewältigung der Multilingualität
- interaktives CLIR
- fachbezogenes CLIR: Amaryllis und GIRT

4 Forschungsfragen des CLIR und der Evaluation

Wesentliche Forschungsfragen im Bereich CLIR sind unter anderem folgende:

- die Kombination geeigneter Methoden zur Behandlung der Multilingualität

Wie die Ergebnisse der Ad-Hoc-Tasks von TREC zeigten, lassen sich mit der Kombination verschiedener Varianten eines Retrievalsystems oder mit der Kombination verschiedener Retrievalsysteme durchaus Steigerungen der Leistung erzielen. Dazu liegt auch der Vorschlag zur Kombination im Rahmen des MIMOR-Modells vor. Jedoch sind diese Kombinationen nicht beliebig, sondern von den spezifischen Eigenschaften der Komponenten oder Systeme abhängig, die kombiniert werden. Hier ist immer noch ein weites Feld empirischer und systematischer Forschung.

- die Integration der Ergebnisse in eine Ergebnisliste (Merging)

Das Zusammenführen der gerankten Ergebnisse aus verschiedenen Datenbeständen (in verschiedenen Sprachen) ist kein triviales Problem. Die Spanne der vorhandenen Ansätze reicht von ganz mechanischen (und daher auch theoretisch völlig unbefriedigenden, aber in der Praxis durchaus erfolgversprechenden) Verfahren zu ausgeklügelten Überlegungen hinsichtlich der Normierung und Integration unterschiedlicher RSV der Einzelergebnisse in das endgültige Ranking.

- die Präsentation der Ergebnisse für den Benutzer (einschließlich multilingualer Hilfsmittel)

Hier lassen sich verschiedene Verfahren der Ergebnispräsentation wie automatische Zusammenfassungen und deren Übersetzung oder die Wiedergabe

wichtiger Sätze oder Phrasen in der Anfragesprache des Benutzers denken. Aber auch die Visualisierung von Ergebnismengen fiele in diese Kategorie.

- die Interaktion von System und Benutzer

Die Interaktivität von Systemen bietet einen weiteren Gesichtspunkt der nach wie vor intensiver Forschung bedarf, hierbei stellt die Multilingualität eine zusätzliche Problemebene dar.

- die Weiterentwicklung der Evaluationsmethodik

Die Methodologie der Evaluation ist sowohl in Bezug auf die Poolingmethode als auch im Hinblick auf neue Aufgabenstellungen des CLIR (Websearch, Question & Answering, CLIR im Multimedia-Kontext usw.) weiterzuentwickeln und weiterhin kritisch zu durchleuchten.

- die vergleichende Analyse der vorgelegten Resultate aus den vergangenen Tests

Nunmehr liegen die Ergebnisse einer Reihe von Evaluationskampagnen vor, die eine Vielzahl von Daten lieferten, die der weiteren Analyse harren. Insbesondere die vergleichende Analyse ähnlicher Verfahren hinsichtlich ihrer Spezifika und der Wirkungen bestimmter Ausprägung der Systeme.

5 Ausblick auf die weitere Entwicklung von CLEF

Gegenwärtig wird die weitere Entwicklung des CLEF nach Ablauf der bestehenden Projektförderung durch die EU-Kommission diskutiert. Für die Planungen des 6. Forschungsrahmenprogramms der EU wurde eine Interessenbekundung für ein Exzellenznetzwerk abgegeben, an der sich eine Vielzahl an Instituten und Firmen aus ganz Europa beteiligt hat. Generell geht es um die Einbeziehung weiterer Sprachen in die Hauptaufgabenstellung des multilingualen Retrieval, zunächst werden wohl Niederländisch, Schwedisch und Finnisch hinzukommen. Wichtig ist auch die Erweiterung um Sprachen aus unterschiedlichen Sprachfamilien wie z.B. Russisch (slawische Sprache mit kyrillischen Zeichen), eine weitere slawische Sprache mit lateinischen Zeichen (aber Ligaturen), eine weniger gesprochene bzw. regionale Sprache wie Baskisch oder Rätoromanisch.

6 Literaturangaben

- Braschler, M., 2002: CLEF 2002 - Overview of Results. In: C. Peters (Hrsg.) Workshop of the Cross-Language Evaluation Forum, CLEF 2002, Rome, Italy, September 2002, Sophia-Antipolis: ERCIM (erscheint)
- Braschler, M., 2002: CLEF 2001 - Overview of Results. In M. Braschler, J. Gonzalo, M. Kluck, C. Peters (Hrsg.), Evaluation of Cross-Language Information Retrieval Systems. Second Workshop of the Cross-Language Evaluation Forum, CLEF 2001. Darmstadt, Germany, September 2001. Revised papers, Berlin et. al: Springer (LNCS 2406)
- Braschler, M., 2001: CLEF 2000 - Overview of Results. In: C. Peters (Hrsg.), Cross Language Information Retrieval and Evaluation, Workshop of the Cross-Language Evaluation Forum, CLEF 2000, Lisbon, Portugal, September 2000, Revised Papers. Berlin et al.: Springer (LNCS 2069), 89-101
- Harman, D., M. Braschler, M. Hess, M. Kluck, C. Peters, P. Schäuble, and P. Sheridan, 2001. CLIR Evaluation at TREC. In: C. Peters (Hrsg.), Cross Language Information Retrieval and Evaluation, Workshop of the Cross-Language Evaluation Forum, CLEF 2000, Lisbon, Portugal, September 2000, Revised Papers. Berlin et al.: Springer (LNCS 2069), 7-23
- Kluck, M., T. Mandl and C. Womser-Hacker 2002. Cross-Language Evaluation Forum (CLEF) – Europäische Initiative zur Bewertung sprachübergreifender Retrievalverfahren. In: Information Wissenschaft & Praxis 53:82-89
- Kluck, M., C. Womser-Hacker 2002: Inside the Evaluation Process of the Cross-Language Evaluation Forum (CLEF): Issues of Multilingual Topic Creation and Multilingual Relevance Assessment. In: M. G. Rodríguez, C. P. S. Araujo (Hrsg.): Proceedings of the Third International Conference on Language Resources and Evaluation, LREC 2002, Las Palmas de Gran Canaria 29-31 May 2002, Paris: ELRA, 573-576 preprint: <http://www.educat.hu-berlin.de/~kluck/lrec-paper-final.doc> oder <http://www.educat.hu-berlin.de/~kluck/lrec-paper-final.pdf>
- Voorhees, E., 2001. Philosophy of IR Evaluation. In: C. Peters (Hrsg.): Results of the CLEF 2001 Cross-Language System Evaluation Campaign, Working Notes for the CLEF 2001 Workshop, 3 September, Darmstadt, Germany, Sophia-Antipolis: ERCIM, 257-260
- Voorhees, E., 2000. Variations in Relevance Judgements and the Measurement of Retrieval Effectiveness. Information Processing & Management 36:679-716
- Voorhees, E. and D. Harman, 2001. Overview of the Ninth Text Retrieval Conference (TREC-9). In: E. Voorhees, D. Harman (Hrsg.) The Ninth Text REtrieval Conference (TREC 9). Gaithersburg: NIST, 1-14, http://trec.nist.gov/pubs/trec9/t9_proceedings.html
- Womser-Hacker, Ch. (1996), Das MIMOR-Modell. Mehrfachindexierung zur dynamischen Methoden-Objekt-Relationierung im Information Retrieval. Habilitationsschrift, Universität Regensburg.



Evaluierung von Internet-Verzeichnisdiensten mit Methoden des Web-Mining

Thomas Mandl

Informationswissenschaft
Universität Hildesheim
Marienburger Platz 22
31141 Hildesheim
mandl@uni-hildesheim.de

Zusammenfassung

Dieser Artikel gibt einen Überblick über die Möglichkeiten und Chancen der Bewertung der Popularität von Internet-Angeboten durch die Analyse der Verlinkungsstruktur. Insbesondere die auf eine Seite verweisenden Hypertext-Verbindungen werden immer häufiger als wichtiger Maßstab für die Autorität und sogar die Qualität von Seiten herangezogen. Eine Methode für die Adaption dieses seiten-orientierten Maßes für komplexere Internet-Angebote wird vorgestellt. Als Untersuchungsbeispiele wurden exemplarisch deutschsprachige Web-Verzeichnisdienste gewählt. Dabei konnte u.a. gezeigt werden, dass die Wahrscheinlichkeit für Hypertext-Verbindungen auf Verzeichnisdienste mit zunehmender hierarchischer Tiefe der Seiten extrem stark abnimmt und dass die Anzahl der Links auf eine Seite eines Verzeichnisdienstes nicht mit der Anzahl der Links auf die dort genannten Seiten konform geht.

Abstract

This paper provides an overview on the possibilities of the evaluation of the popularity of web sites by analyzing link structure. Especially back links are more and more considered as an important indicator for the authority and even the quality of pages. The most important algorithms for these tasks are presented. Furthermore, this page oriented method is extended for complex sites. The possibilities of web structure mining are explored by analyzing German web catalogues. It could be shown that the probability for a link to point to a page in a catalogue decreases drastically when the page is low in the hierarchy and that the number of links to a page in an internet catalogue does not correlate with the number of back links of the sites mentioned.



1 Einleitung

Die Popularität der Suchmaschine Google und des darin verwendeten Page-Rank Algorithmus, der die Verlinkung der Hypertext-Struktur des Internets als Ergänzung für das Ranking der Ergebnis-Dokumente heranzieht, hat zu einem breiten Interesse an der Bewertung von Internet-Angeboten durch Link-Analysen geführt. Die Anzahl von auf eine Internet-Seite verweisenden Hypertext-Verbindungen (*back links* oder *in links*) gilt als Maß für die Autorität und Popularität dieser Seite. Die unkritische Verwendung der einfachen Maßzahlen wirft jedoch zahlreiche Fragen auf. Dieser Beitrag will einige dieser Probleme im Kontext einer Evaluierung von Web-Sites vorstellen.

Da die Link-Analysen meist im Rahmen von Schlagwort-Suchen genutzt werden, wurden hier bewusst Angebote gewählt, die das Browsing als primäre Suchstrategie nutzen. Das Verfolgen von Hyperlinks ist eine sehr wichtige Strategie für die Orientierung (MARCHIONINI 1995), die häufig unterschätzt wird. Eine Untersuchung von MAHOUI & CUNNINGHAM 2001 zeigt, dass in der untersuchten digitalen Bibliothek lediglich ca. 50% der Benutzer eine Suchanfrage stellen. Vermutlich wünschen viele der übrigen Besucher die Möglichkeit, ihr Informationsbedürfnis durch effiziente Browsing-Angebote zu lösen.

Browsing im Internet basiert auf Verknüpfungen, die entweder vage Ähnlichkeitsbeziehungen widerspiegeln wie etwa in zweidimensionalen Karten (cf. EIBL & MANDL 2002) oder fest definierte Beziehungen zeigen, die wie etwa in Verzeichnisdiensten (*web directories*, *internet catalogues*). Die Bedeutung dieser Dienste und ihrer Bewertung steigt, da sie inzwischen auch automatisiert erstellt werden (cf. z.B. FROMMHOLZ 2001). Dazu werden Verfahren der Text-Kategorisierung (MANDL 2001) auf hierarchische Strukturen übertragen.

Die Verzeichnisdienste zeichnen sich vor allem durch ihre hierarchische Struktur aus. Ansätze zu ihrer Bewertung lassen sich demnach auch auf zahlreiche andere hierarchische Web-Angebote übertragen. Hier haben sich automatische Verfahren für die Evaluierung noch wenig etabliert. RITTBERGER 2001 stellt etwa eine umfassende Methodik vor, welche die intellektuelle Bewertung ermöglicht.

2 Web-Mining

Web-Mining nennt sich eine Forschungsrichtung, die auf die Gewinnung von Wissen aus den großen Datenmengen im Internet abzielt. Dabei geht es nicht

um das Retrieval einzelner Seiten mit einem geringen Funktionsumfang, sondern um die Zusammenfassung sehr vieler Daten zu neuem Wissen, das so explizit nicht gespeichert ist. Der Begriff entstand als Analogie zum Data-Mining.

Data-Mining war eine Konsequenz der großen Datenmengen, die mit fallenden Preisen für Massenspeicher in immer mehr Anwendungsbereichen vorlagen: "Now that we have gathered so much data, what do we do with it?" (FAYYAD & UTHURUSAMY 1996:24). Die Definition eines Algorithmus erinnert noch stark an das maschinelle Lernen: "... Any algorithm that enumerates patterns from, or fits models to, data is a *data mining algorithm*" (FAYYAD 1997:5). Allerdings soll Data-Mining den gesamten Prozess der Sammlung und Pflege der Daten angefangen, über ihre Auswertung bis hin zu der Anwendung der Ergebnisse umfassen.

Web-Mining überträgt Data-Mining auf Anwendungen im Internet und passt insbesondere die Systeme zum Sammeln der Daten auf die Gegebenheiten im Web an. In diesem Kontext wurde bereits der Begriff *Web-Intelligence* geschaffen, der auch schon Titel für eine wissenschaftliche Konferenz war¹. Folgende Teilaspekte von Web-Mining sind nach WALTHER 2001 relevant:

- Web-Content-Mining
- Web-Structure-Mining
- Web-Usage-Mining

Diese Teilaspekte werden im Folgenden kurz erläutert.

2.1 Web-Content-Mining

Content-Mining versucht im weitesten Sinne Inhalte von Web-Seiten auszunutzen und durch das Zusammenfassen von großen Mengen von Inhalten zu neuem Wissen zu gelangen.

Dazu zählt etwa das Extrahieren von mehrsprachigen Parallel-Korpora aus dem Web für die maschinelle Übersetzung oder multilinguales Information Retrieval. Dabei werden anhand heuristischer Regeln, die z.B. auf Web-Servern nach Verzeichnissen mit gleichnamigen Dateien suchen, deren Verzeichnisnamen auf Sprachversionen hinweisen (cf. z.B. NIE ET AL. 2001). Die von Menschen geleistete Übersetzungsarbeit für diese konkreten Texte wird dann in einem Transfer-Prozess ausgenutzt, um durch die Parallelisierung von

¹ <http://kis.maebashi-it.ac.jp/wi01>.

Texten in verschiedenen Sprachen eine Übersetzung vorzunehmen, die auf den Kontexten basiert, in denen ein Wort vorkommt (cf. MANDL 2001).

Andere Ansätze versuchen anhand von Trends in Publikationen auf sinnvolle Forschungsrichtungen zu schließen (PIROLI ET AL. 1996).

2.2 Web-Structure Mining

Das Ziel des Structure-Mining besteht darin, Wissen aus der Struktur von Hypertexten zu gewinnen. Struktur kann sich dabei auf die HTML Struktur einzelner Dokumente beziehen, auf die Struktur von Sites und auf die Verlinkungsstruktur des gesamten Internets oder von Ausschnitten des Internet wie etwa den Resultaten von Suchanfragen. Meist sollen die Struktur-Informationen die Suche im Web verbessern (cf. z.B. WEISS ET AL. 1996, BEKAVAC 1999, KRUSCHWITZ 2001). Diese Verfahren beschreibt das nächste Kapitel ausführlich.

2.3 Web-Usage-Mining

Beim Web-Usage-Mining geht es um die Interpretation von sehr vielen Benutzeraktionen im Internet. Daraus lassen sich Folgerungen über die Angemessenheit und Gebrauchstauglichkeit von Angeboten ziehen.

Als Quelle dienen meist die Log-Dateien einzelner Server oder großer Proxy-Server. Ebenso können lokale Bookmark-Dateien, lokale History-Dateien von Web-Clients oder spezifische Client-Log-Dateien genutzt werden.

Die Log-File-Analyse basiert i.d.R. auf dem Server-Log eines Anbieters und kann für die Analyse und Optimierung der Seiten dieses Anbieters eingesetzt werden (cf. z.B. BARTEL 2002). Ein Vergleich mit anderen Sites für ähnliche Anwendungen (wie etwa der Vergleich zweier Online-Geschäfte für Bücher) ist so nicht möglich. Dazu wäre eine globale Perspektive erforderlich, aus der sich die Interaktionen vieler Benutzer mit vielen Internet-Angeboten ansehen lässt. Dies würde am ehesten durch die Proxy-Server großer Internet Service Provider (ISP) erreicht, die allerdings dem Datenschutz unterliegen.

Die Analyse von Nutzungsdaten aus dem Internet gilt auch als eine große Chance für die empirische Forschung zur Mensch-Maschine-Interaktion. Die Evaluierung von Benutzungsschnittstellen ist extrem wichtig, da in einer derart komplexen und noch nicht vollständig erforschten Domäne wie der Mensch-Maschine-Interaktion rein analytische Aussagen zur Qualität einer

Software hinsichtlich der Interaktion kaum getroffen werden können. Um aussagekräftige Ergebnisse zu gewinnen, müssen Benutzertests durchgeführt werden. Dazu gehört die Beobachtung und Befragung einiger repräsentativer Testbenutzer während der Erledigung von Testaufgaben. Alle Interaktionsschritte sollten aufgezeichnet und der Benutzer gefilmt werden. Solche Untersuchungen erfordern den Einsatz von vielen Ressourcen und sind sehr teuer.

Dagegen entstehen die Daten in Log-Files durch tatsächliche reale Aktionen von Internetnutzern. Verfahren für ihre Auswertung gelten daher als „‘discount‘ usability assessment methods“ (SULLIVAN 1997:2). Allerdings enthalten die Log-Files nur sehr wenig Information und sind daher keineswegs so aussagekräftig wie Benutzertests. So zeichnet der Server lediglich den Abruf einer Datei auf, die Interaktion mit Web-Seiten geschieht lediglich auf Seiten des Clients, so dass darüber keine Informationen vorliegen. Darüber hinaus sind die Daten der Log-Dateien mit großen Unsicherheiten behaftet. Weder lässt sich ein Benutzer eindeutig identifizieren, noch erreichen alle Seitenaufrufe den Server, da Dateien teils aus dem lokalen Cache oder von Proxy-Servern geladen werden.

3 Analysen der Verlinkungsstruktur

Im Web-Structure-Mining dominieren Analysen der Hypertext-Verknüpfungen (für einen Überblick cf. BAEZA-YATES & RIBEIRO-NETO 1999:380f.). Die Grundideen dieser Analysen der Verlinkungsstruktur stammen aus der Biblio- oder Szionometrie, die das Netzwerk der wissenschaftlichen Zitate analysieren und darauf abzielen, Publikationen aufgrund der Häufigkeit der auf sie verweisenden Zitate zu bewerten. Komplexere Maße betrachten z.B. die Stellung eines Autors im Diskurs-Netzwerk und berücksichtigen über die Zitate hinaus die institutionelle Zugehörigkeit und Ko-Autorenschaft (cf. MUTSCHKE 2001). Andere komplexere Analysen errechnen aus den Häufigkeiten von Zitaten, Maßzahlen für das Renommee von Zeitschriften, Tagungen oder Fachbereichen (cf. z.B. SCHLÖGL 2000).

Die technischen Möglichkeiten der online Verfügbarkeit von wissenschaftlicher Literatur führt dazu, dass bibliometrische Analysen heute Teil von kostenlos zugänglichen digitalen Bibliotheken sind².

² Entsprechende Kennzahlen sind z.B. in Daffodil (<http://www.daffodil.de>) und Research-Index (<http://citeseer.nj.nec.com/cs>) integriert.

3.1 Autoritätsmaße

Die Suchmaschine Google kann als umfangreichste und erfolgreichste Implementierung einer automatischen Analyse der Verlinkungsstruktur gelten. Der in Google wirkende PageRank Algorithmus (PAGE ET AL. 1998) zählt nicht nur die Links auf eine Seite. Zunächst erhalten alle Seiten das gleiche Gewicht als Verteiler. Das bedeutet, dass die Autorität, die eine verweisende Seite vergeben kann, an der Anzahl der ausgehenden Links relativiert wird. Darüber hinaus wird der Einfluss einer Seite auch mit deren Autorität relativiert. Je größer die Autorität einer Seite ist, desto höheres Gewicht haben die von ihr ausgehenden Links.

Der Algorithmus arbeitet iterativ. Zunächst werden alle Seiten mit dem gleichen Autoritätswert initialisiert und dann berechnet der erste Schritt die neue Autorität aller Seiten aus der Verlinkung. Dabei ergeben sich neue Autoritätswerte, so dass alle Werte nun erneut berechnet werden und das Ergebnis die angestrebte Autorität besser wiedergibt. Laut den Autoren konvergiert der Algorithmus nach einer Anzahl von Schritten (cf. PAGE ET AL. 1998), d.h. bei einem weiteren Berechnungsschritt verändern sich die Autoritätswerte kaum mehr. Der Autoritätswert wird für das Berechnen des Rankings der Dokumente nach einer Anfrage benutzt (cf. PAGE ET AL. 1998).

Der PageRank-Algorithmus benutzt sinnvolle Annahmen, um die Autorität zu berechnen. Allerdings beschränkt er sich auf die Ebene der Seite und berechnet z.B. per se nicht die Autorität einer gesamten Site.

Laut PAGE ET AL. 1998 gibt der PageRank-Algorithmus die Wahrscheinlichkeit wieder, mit der ein Surfer auf eine Seite trifft, wenn er Hypertext-Verbindungen verfolgt und nie auf eine bereits besuchte Seite zurückführt. In der Suchmaschine Google wird der PageRank-Wert einer Seite mit der System-Relevanz kombiniert. Die Wahrscheinlichkeit des Treffens auf eine Seite beim Browsen wird also auf die Suche übertragen.

3.2 Autorität bei der Vermittlung

Der PageRank-Algorithmus kann als eine einfachere Version des Kleinberg-Algorithmus betrachtet werden, die aber für wesentlich größere Mengen von Seiten geeignet ist. Dieser zielt ebenfalls auf Autorität ab und berücksichtigt nur die Verbindungsstruktur zwischen einer Menge von Seiten. Kleinberg führt zwei Rollen ein, um die Autorität zu bewerten (KLEINBERG 1998). Er spricht von hubs und authorities und weist jeder Web-Seite ein Gewicht für

beide Rollen zu. Der hub entspricht einem Mittelpunkt, Verteiler oder Informationsvermittler, dessen Aufgabe im Wesentlichen in der Bereitstellung von Verbindungen zu anderen Seiten besteht. Dahinter steht die Vorstellung eines clearinghouses oder in der Wissenschaft der eines guten Überblicksartikels. Die authorities dagegen enthalten die eigentliche Information in unterschiedlicher Qualität.

Im Gegensatz zum PageRank-Algorithmus findet das Verfahren von Kleinberg nur Anwendung auf eine Menge von ca. 5.000 bis 10.000 Seiten, die aus einer Suchanfrage ermittelt werden. Die besten Suchergebnisse eines Suchdienstes werden analysiert und die enthaltenen Verbindungen extrahiert. Die entsprechenden Seiten gelangen bis zu einer bestimmten Tiefe in den Datenbestand. Die Verbindungen innerhalb dieser Menge werden nun iterativ analysiert. Jede Seite besitzt sowohl ein Gewicht als hub als auch als authority, die in jedem Durchlauf modifiziert werden. Die Autorität einer Seite steigt mit der Anzahl der ankommenden Verbindungen. Diese Zahl wird aber an dem hub-Gewicht der Ausgangsseite relativiert. Nur die Links von guten Verteilerseiten wirken sich somit stark auf die Autorität einer Seite und damit auf das authority-Gewicht aus. Ebenso unterliegt das hub-Gewicht einer Veränderung, die von der Autorität der Zielseiten abhängt. Auf je bessere Seiten der Verteiler verweist desto besser ist er und desto stärker steigt sein hub-Gewicht. Ziel ist die Identifikation der Seiten mit der höchsten Autorität innerhalb der Untermenge.

Der Algorithmus birgt die Gefahr der weiten thematischen Entfernung durch die Integration weiterer Seiten neben dem eigentlichen Suchergebnis. Diese können von dem Thema, das mit der Suchanfrage verbunden ist, schon weit entfernt liegen (*topic-drift*).

Weitere Algorithmen wurden von GIBSON ET AL. 1998 und LEMPEL & MORAN 2000 vorgeschlagen. MEGHABGHAB 2002 untersucht die Algorithmen für Linkanalysen aus dem Blickwinkel der Grafentheorie und Matrizenalgebra und leitet einige formale Eigenschaften der Berechnungsmethoden ab.

3.3 Evaluierung

Inwieweit die Autoritätsmaße wie PageRank für die Benutzer Vorteile bringen, müssen Evaluierungen zeigen. Für das Information Retrieval gibt es Hinweise, dass sich PageRank bei großen Kollektionen positiv auf die Retrieval-Qualität auswirkt.

Im Rahmen der Evaluierungsstudie TREC (Text Retrieval Conference, cf. VOORHEES & HARMAN 2001) wird die Leistung von Information Retrieval Systemen hinsichtlich der Fähigkeit gemessen, thematisch relevante Dokumente zu identifizieren. Die Bedeutung von TREC liegt in dem großen Umfang der Daten und der erreichten Vergleichbarkeit zwischen Systemen. TREC stellt sich in seiner zehnjährigen Geschichte auch zahlreicher Kritik und neuen Anforderungen. Kritik richtete sich u.a. gegen die Methodik, welche den Benutzer vernachlässigt. Als eine Konsequenz wurde ein *interactive track* eingeführt, welcher die Unterstützung des Benutzers durch die Benutzeroberfläche untersucht.

Die neuen Anforderungen ergaben sich zu einem großen Teil aus der Realität im Internet und der dortigen Verfügbarkeit von Information Retrieval Systemen. Als Konsequenz wurden sehr große Kollektionen und kurze Anfragen betont und Mehrsprachigkeit und multimediale Objekte rückten in den Fokus.

Die Bedeutung des Internet führte zur Einführung des Web-Track (cf. HAWKING 2001), bei dem nicht Zeitungstexte die Grundlage bilden sondern Internet-Dokumente. Um das übliche TREC Prozedere beibehalten zu können und den Systemen eine feste Datenmenge und Übungszeit zu bieten, speichern die Veranstalter eine Momentaufnahme eines Teils der im Internet angebotenen Daten. Davon liegen zwei verschieden große Versionen vor³, welche hinreichend groß sind, um die Wirksamkeit von linkbasierten Verfahren zu testen. Bei der intellektuellen Überprüfung der Ergebnisse der Suchmaschinen achten die Juroren auf thematische Relevanz der Dokumente zu den formulierten Topics. Neben den thematischen Suchen wurde als weitere Aufgabe das Finden von *Homepages* eingeführt, für die als Anfrage z.B. der Name einer Institution vorlag.

PageRank wurde von der Universität Neuchatel in TREC eingesetzt (cf. SAVOY & RASOLOFO 2000). Die Ergebnisse des Web-Track in TREC mit PageRank und anderen Verfahren weisen darauf hin, dass die Berücksichtigung von Hypertext-Verknüpfungen die Ergebnisse des Retrievals verbessern kann. Dies gilt bereits bei den in TREC verwendeten Momentaufnahmen, die natürlich bei weitem nicht das gesamte Internet umfassen. Damit werden weder alle Links auf die in dem Sample enthaltenen Seiten erfasst, noch können alle in dem Sample vorkommenden Verknüpfungen benutzt werden, weil viele von ihnen auf Seiten außerhalb verweisen. Die Verbesserung der Ergebnisse konnte bisher nur für die Suche nach *Homepages* und damit für Suchen nach

³ Die kleine Momentaufnahme besteht aus 1,7 Millionen Seiten (10 Gigabyte), während der große *snapshot* 18,5 Millionen Seiten (100 Gigabyte) umfasst (HAWKING 2001:1).

einer konkreten Seite nachgewiesen werden, während für die thematisch orientierten Suchen keinerlei Vorteile gemessen wurden (HAWKING 2001:10).

3.4 Nachteile

Die Nachteile von Autoritätsmaßen wie dem PageRank-Algorithmus liegen auf der Hand:

- Web-Seiten werden unabhängig ihres Inhalts und Kontexts bewertet. Ebenso wie Wissenschaften unterschiedliches Zitationsverhalten aufweisen, ist davon auszugehen, dass in verschiedenen Internet-Dokument-Typen unterschiedliche Verlinkungsneigung herrscht.
- Kritiker von bibliometrischen Analysen bemängeln, dass die Dynamik und Pragmatik des wissenschaftlichen Publizierens durch einfache Kennzahlen nicht hinreichend abgebildet wird (cf. z.B. FRÖHLICH 2000).
- Die Algorithmen können manipuliert werden. Dies geschieht sicher in hohem Maße, da ein erhebliches wirtschaftliches Interesse daran besteht, eigene Seiten bei vielen Internet-Suchen an vorderen Ranking-Positionen zu sehen.
- Die Algorithmen beruhen auf plausiblen Annahmen, berücksichtigen aber nicht die Sicht der Benutzer.
- Die Maße beziehen sich auf einzelne Seiten und können nicht direkt auf komplexere Angebote übertragen werden.

4 Entwicklung der Evaluierungsmethode

Um die Argumente für und wider besser abwägen zu können, wird im Folgenden versucht, diese Verfahren auf sogenannte Verzeichnisdienste im Internet (Web-Kataloge) anzuwenden. Die in dieser Studie entwickelte Methode zielt darauf ab, mit den Verfahren der Linkstrukturanalyse zu Erkenntnissen über die Benutzung der Verzeichnisdienste zu gelangen. Dabei muss das sich an Einzelseiten orientierte Maß In-Links auf komplexere Hypertext-Strukturen bezogen werden. Die Struktur der untersuchten Dienste ist durchweg hierarchisch.

4.1 Fragestellungen

Die Untersuchung soll vor allem folgende Aspekte untersuchen:

- Verweisen Links auf Verzeichnisdiensten eher auf Seiten, die in der Hierarchie weit oben stehen oder auf Unterkategorien zu spezifischen Themen?

- Zeichnen sich Unterschiede zwischen Verzeichnisdiensten für die Autorität ähnlicher Themen ab?

Die ethische Problematik des sog. *deep linking* dürfte hierbei kaum eine Rolle spielen. Von diesem erst jüngst wieder in den Schlagzeilen erscheinenden Phänomen spricht man, wenn Hyperlinks nicht auf die Homepage eines Web-Auftritts verweisen, sondern auf tiefer gelegene Seiten. Problematisch erscheinen solche Links, wenn der entsprechende Anbieter ein Geschäftsmodell verfolgt, nach dem er auf der Homepage Werbung platziert und auf den tieferen Seiten einen informationellen Mehrwert bietet (cf. SPINELLO 2001). Die Internetkataloge werben wenn überhaupt auch auf hierarchisch tieferen Seiten.

4.2 Systemarchitektur

Für diese Untersuchung wurde ein spezielles Web-Mining-System in JAVA 1.4 entwickelt, das auf mehreren frei im Internet verfügbaren Komponenten aufbaut. Die Seiten werden mit einem Roboter aus dem Netz übertragen. Dazu wurde ein Opensource-Roboter an die Bedürfnisse der Anwendung angepasst⁴. Die HTML-Seiten werden mit dem Tidy-Parser von W3C analysiert und in ein *Document Object Model (DOM)* Repräsentation überführt⁵. Anhand von Heuristiken wurden Regeln für jeden Verzeichnisdienst entworfen, die beschreiben, wie das System innerhalb der DOM-Repräsentation u.a. Titel, Hierarchie-Level, Unterkategorien und die eigentlichen Verweise auf andere Web-Angebote erkennt.

```
<?xml version="1.0" encoding="UTF-8"?>
<html><!-- ysx:2025869239 --><!-- 020513:0014 --><head>
  <meta content="HTML Tidy, see www.w3.org" name="generator"/>
  <title>Yahoo! Nachschlagewerke&gt;Statistiken</title>
  <base href="http://de.dir.yahoo.com/Nachschlagewerke/Statistiken/" />
</head>
<body bgcolor="ffffff">
  <table cellpadding="0" cellspacing="0" width="100%" border="0">
    <tr>
      <td width="1%">
        <a href="http://de.yahoo.com">
          
        </a>
      </td>
      <td>
        <table width="100%" cellpadding="0" cellspacing="0" border="0">
          <tr>
```

Abb. 1: Ausschnitt aus der DOM-Repräsentation einer Yahoo-Seite

⁴ <http://www.matuschek.net/software/jobbo/index.html>.

⁵ <http://w3c.org>.

Die Ergebnisse wurden in einer relationalen Datenbank abgelegt und mit einer Tabellenkalkulationssoftware ausgewertet.

Um die eingehenden Verbindungen einer Web-Seite zu erhalten, ist theoretisch eine vollständige Analyse des Internets nötig. Dies erfordert einen immensen Aufwand an Software und ist nicht bei jeder Untersuchung zu leisten. Die Crawler von Suchmaschinen durchsuchen das Web ohnehin und einige erlauben auch, die Links auf eine Seite abzufragen. Google und Altavista gehören dazu. In diesem Experiment wurde mit der Adresse der jeweiligen Seite eine gültige Abfrage in der Syntax der beiden Sprachen gestellt und die zurückgelieferte Ergebnisseite nach der entsprechenden Zahl durchsucht. Verbindungen aus dem Verzeichnisdienst selbst wie sie für die Interaktion unvermeidlich sind, wurden nicht eliminiert.

Experimente im Web-Mining leiden noch unter anderen Unwägbarkeiten und Schwierigkeiten, die zu einer gewissen Vagheit bei den Ergebnissen führen. Bricht etwa während der langen Zeit des Retrievals die Netzverbindung kurzzeitig ab und überschreitet die Ausfallzeit einen bestimmten timeout, dann werden möglicherweise einige Seiten nicht erfasst. Dies gilt auch für die Abfrage der In-Links bei Google und Altavista, die durchaus auch kurzzeitig nicht erreichbar sind.

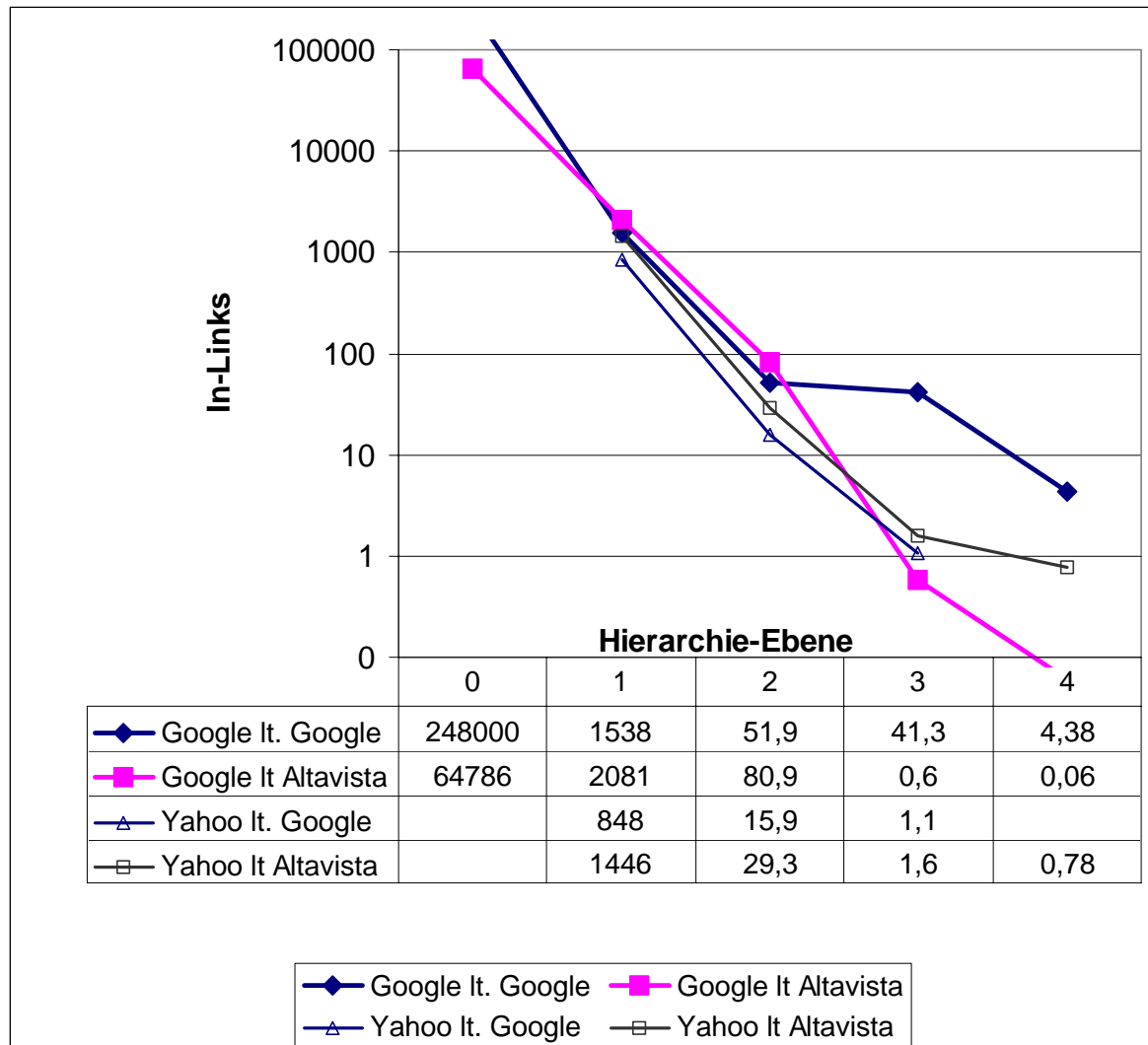
5 Ergebnisse der Evaluierung deutschsprachiger Internet-Verzeichnisdienste

Die Bewertung von Verzeichnisdiensten anhand von Linkanalysen bietet einen Einblick in die Benutzung von Verzeichnisdiensten allgemein. Die Analyse zeigt an einigen Parametern, wann Verbindungen auf Seiten in Verzeichnisdiensten gelegt werden.

Die Untersuchung konnte vor allem beweisen, dass die Zahl der Verbindungen auf einen Verzeichnisdienst sehr stark abnimmt, je tiefer die Seite in der hierarchischen Struktur steht. Dazu wurde der Dienst Google auf mehreren Ebenen untersucht. Trägt man die Abhängigkeit in einer logarithmischen Skala ein, dann ergibt sich ein fast linearer Zusammenhang (siehe Abb. 2). Die Ebene Null entspricht dabei der Einstiegsseite⁶, während die folgende Adresse ein Beispiel für eine Seite der vierten Ebene darstellt: Computer/Software/

⁶ <http://directory.google.com/Top/World/Deutsch/>.

Spezialsoftware/Sport⁷. Analysiert wurden ca. 3000 Seiten des Internet-Katalogs von Google und ca. 500 Seiten der deutschen Version von Yahoo.
Abb. 2: Zahl der durchschnittlichen In-Links pro Ebene



Die extrem starke Abnahme der eingehenden Links von meist mehr als den Faktor zehn weisen sowohl die Seiten von Google als auch die von Yahoo auf. Ähnliche Beziehungen ergaben sich auch bei anderen Analysen eingehender Links von ADAMIC & HUBERMANN 2001. Darin konnten die Autoren zeigen, dass die Zahl der Seiten mit vielen eingehenden Verbindungen sehr stark abnimmt bis hin zu Seiten ohne In-Links. Auch dabei ergab sich auf einer logarithmischen Skala ein linearer Zusammenhang.

⁷ <http://directory.google.com/Top/World/Deutsch/Computer/Software/Spezialsoftware/Sport>.

Obwohl die Ergebnisse von Google und Altavista für die Abfrage der In-Links oft sehr stark voneinander abweichen, gilt der oben vorgestellte Zusammenhang bei beiden Suchmaschinen.

Die Abhängigkeit, die sich hier für eine Untermenge der Google- und Yahoo-Seiten ergibt, gilt möglicherweise auch für andere hierarchisch aufgebaute Seiten.

Die Autoren von Internet-Seiten verweisen also grundsätzlich eher auf generelle Seiten in Verzeichnisdiensten und weniger auf Seiten zu spezifischen Themen. Offensichtlich wollen sie dem Benutzer die Suche nach für ihn interessanten Angeboten durch Browsing selbst überlassen. Möglicherweise bewerten die Autoren Seiten der Verzeichniseinträge als gut geeignet für thematisch breite Informationsbedarfe oder für Informationsbedarfe, die noch vage sind und bei denen der Benutzer sich zunächst einen Überblick verschaffen will. Eine Interpretation, dass die redaktionelle Arbeit für die Dienste nicht gewürdigt wird, geht sicher zu weit.

Bei der Anzahl der Einträge pro Seite und der Anzahl der Unterkategorien (mit Querverweisen) ergaben sich keine interessanten Aspekte.

5.1 Vergleich von Verzeichnisdiensten

Ein Vergleich von Verzeichnisdiensten aufgrund von In-Links muss berücksichtigen, dass auf den unteren Ebenen grundsätzlich weniger Verbindungen zu erwarten sind. Im Folgenden werden die Startseiten mehrerer Dienste verglichen. Dabei ist allerdings ist die Unsicherheit größer, die sich aus der Abfrage zweier unterschiedlicher Suchmaschinen nach den eingehenden Verbindungen ergibt. Folgende Tabelle zeigt die eingehenden Links für einige deutschsprachige Verzeichnisdienste bzw. die deutschsprachigen Unterkategorien global angelegter Verzeichnisdienste.

Verzeichnis- dienst	Allesklar	Altavista	Dino	dmoz.org	Google	Web.de
Lt. Google	2390	374	3610	19000	14700	22
Lt. Altavista	51581	1905	32999	1886	27110	14946

Tabelle: Links zu den Startseiten einiger deutscher Verzeichnisdienste

Die Unterschiede zwischen der Anzahl von ankommenden Verbindungen zwischen Google und Altavista sind teilweise sehr hoch, so dass die Aussage-

kraft dieser Zahlen nicht sehr groß ist. Die geringen Werte von Web.de und Altavista bei Google erklären sich daher, dass weitaus mehr Links auf die Suchseiten zeigen als auf die Verzeichnis-Seiten.

Eine weitere Vergleichsmöglichkeit besteht in der Analyse der Seiten zu ähnlichen Themen. Eine Automatisierung ist hier schwierig, da nicht in allen Fällen die gleichen Beschreibungen verwendet werden. Dazu wurden zwei Seiten auf Ebene drei und vier ausgewählt, die in Google in dieser Ebene die höchsten Werte erreichten. Dies waren eine Seite zu Geisteswissenschaften/ Geschichte und eine zu Internet/WWW/Firmen. Diese Seiten ragen aus den sonst auf dieser tiefen Ebene wenig referenzierten Seiten deutlich heraus. In Yahoo hatte eine vergleichbare Seite zur Geschichte auch sehr viele In-Links, ohne dass hier eine vollständige Analyse aller Seiten durchgeführt worden wäre. Dagegen wurde auf eine ähnliche Seite mit Firmen für das WWW sehr viel weniger häufig verlinkt. Dies kann als Hinweis darauf gelten, dass thematisch ähnliche Seiten verschiedener Verzeichnisdienste sehr unterschiedliche Qualität aufweisen.

5.2 In-Links von referenzierten Seiten

Ein wichtiger Maßstab für die Qualität von Verzeichnisdiensten kann neben der Anzahl von Verbindungen, die auf die Seiten des Dienstes verweisen, die Anzahl der Verbindungen für die darin enthaltenen Angebote sein. Korreliert also die Zahl der Verbindungen auf eine Seite eines Verzeichnisdienstes mit der Autorität der Seiten, auf die der Dienst verweist, gemessen an den Verweisen auf diese Seiten? Finden die Redakteure der Dienste solche Web-Angebote, die auch laut Linkanalyse eine hohe Autorität besitzen? Oder unterscheidet sich das menschliche Urteil von den automatisch ermittelten Zahlen?

Um dies näher untersuchen zu können, wurde bei der oben bereits erwähnten Stichprobe von ca. 400 Seiten aus Yahoo neben den Links auf diese Seiten auch die dort verzeichneten Seiten untersucht. Für jeden Eintrag in diesen Yahoo Seiten wurde eine Anfrage sowohl an Google als auch an Altavista geschickt, um die Anzahl der *back links* dieser Angebote zu erhalten. Für jede Yahoo-Seite wurde der Durchschnitt jeweils für die Google- und die Altavista-Ergebnisse gebildet.

Zunächst zeigt sich, dass eine sehr große Streuung besteht. Bei einem Mittelwert von 426 eingehenden Verbindungen weisen die Ergebnisse von Google eine Standardabweichung von 1.346 Verbindungen auf. Altavista liefert als

Mittelwert 79 und mit 209 ebenfalls eine sehr hohe Standardabweichung. Die Autorität und damit die vermutete Qualität der referenzierten Seiten wäre demnach sehr unterschiedlich.

Interessanterweise ergibt sich für die Stichprobe keinerlei Korrelation zwischen den *back links* der Verzeichnisdienst-Seite und den *back links* der dort genannten Web-Angebote. Der Betrag der berechneten Korrelation liegt unter 0,1. Angesichts der obigen Ergebnisse, nach denen auf in der Hierarchie tiefer liegende Seiten kaum verwiesen wird, sollte man aber Seiten auf unterschiedlichen Ebenen getrennt betrachten. Möglicherweise überlagert der in Abbildung 2 skizzierte Effekt der starken Abnahme der Wahrscheinlichkeit für eingehende Verbindungen die gesuchte Korrelation. Daher wurden die Korrelationen auch für die Seiten nur einer Ebene berechnet, allerdings ergab sich auch hier keine nennenswerte Korrelation.

Die menschliche und die automatische aus den Verbindungen abgeleiteten Qualitätsurteile stimmen demnach nicht überein.

Von allen untersuchten Parametern ergab sich lediglich für die Anzahl von Unterkategorien einer Seite eine positive Korrelation mit einem Betrag von mehr als 0,5. Demnach steigt die Wahrscheinlichkeit für Verbindungen auf eine Seite eines Verzeichnisdienstes leicht an, wenn die Seite viele Unterkategorien enthält. Dagegen wirkt sich die Anzahl der enthaltenen Verweise auf externe Angebote weder positiv noch negativ aus. Dies konnte so nicht unbedingt erwartet werden, da die Leistung der Dienste gerade in der Bewertung externer Links liegt.

Für denjenigen, der einen Link auf eine Dienst-Seite legt, scheint der Aspekt des Browsing bei der Benutzung im Vordergrund zu stehen. Möglicherweise sollen dem Benutzer, der den Link verfolgt, noch viele Optionen geboten werden. Zum einen wird also auf Seiten verwiesen, die selbst eher viele weitere Kategorien enthalten, aus denen ausgewählt werden kann. Zum anderen wird auch sehr viel häufiger auf Seiten verwiesen, die hoch in der Hierarchie stehen und von denen aus der Benutzer noch mehrere Ebenen weiter verfolgen kann. Damit weisen die beiden Ergebnisse aus der Linkanalyse in die gleiche Richtung.

Die Gründe für das Setzen von Links auf Seiten eines Verzeichnisdienstes umfassen sicher noch viele weitere Aspekte. Und das Erstellen von Hypertext-Verbindungen kann nur ein Anhaltspunkt für die Benutzung von solchen Diensten bieten. Trotzdem lassen sich mit der Linkanalyse interessante An-

haltspunkte gewinnen. Weitergehende Evaluierungen müssten u.a. folgende Aspekte beachten, um zu genaueren Aussagen zu kommen:

- Wie stark überschneiden sich die Einträge in den Diensten? Ist der Unterschied der gemessenen Autorität der Verzeichnisseiten dadurch gerechtfertigt?
- Finden sich die thematisch ähnlichen Seiten auf der gleichen Ebene? Ansonsten muss berücksichtigt werden, dass eine tiefere Einbettung eine geringe Wahrscheinlichkeit für In-Links mit sich bringt.
- Inwieweit unterscheidet sich die Neigung, Verbindungen zu setzen über Themen hinweg? Ebenso wie die Wissenschaften unterschiedliches Zitierverhalten offenbaren, so könnten auch verschiedene Nutzergruppen unterschiedlich häufig Links setzen. Dies könnte über mehrere Verzeichnisdienste hinweg untersucht werden.

5.3 Unterschiede zwischen den Diensten

Allesklar und Web.de linkten nicht direkt auf die Einträge, sondern auf eine interne Datenstruktur, die auf die Site des Eintrags verweist. Dies spielt für den Benutzer kaum eine Rolle, allerdings nutzen die beiden Verzeichnisdienste Allesklar und Web.de auch kryptische, mit Zahlen kodierte URLs für ihre Seiten. Dies verringert die Benutzerfreundlichkeit für verweisende Autoren und damit die Zahl der potenziellen In-Links. Auch Altavista kommt nicht ohne für den Benutzer kryptische Zeichen in der URL aus, da die Entwickler hier Parameter einsetzen. Entgegen den anderen untersuchten Diensten zeigt Altavista die Einträge vor den Unterkategorien an. Für Benutzer, die sich an andere Verzeichnisdienste gewöhnt haben, ist dies ungewohnt und kann sogar dazu führen, dass die Kategorien gar nicht wahrgenommen werden. Dino lässt dagegen nur Einträge auf der untersten Ebene zu, so dass Einträge von externen Seiten und die Kategorien getrennt bleiben. Bei einer Analyse der Einträge und Unterkategorien muss dies berücksichtigt werden.

6 Ausblick

Die Link-Analyse im Internet wird in diesem Beitrag exemplarisch auf Verzeichnisdienste angewandt, um zu zeigen, wie die seiten-bezogenen Kennzahlen auf größere Einheiten wie Sites übertragen werden können. Die Untersuchung zeigt einige interessante Ergebnisse für die Analyse des Web-Verzeichnisses ebenso wie für die themenspezifischen Analysen. Die Wahrscheinlichkeit eines Links auf Seiten eines Internet-Katalogs nimmt mit der

hierarchischen Tiefe dieser Seiten sehr stark ab. Die Benutzung der Kataloge beschränkt sich für das Setzen von Links auf sehr generelle Seiten.

Vergleicht man die Anzahl der Links auf eine Seite in einem Katalog, so korreliert sie nicht mit der Anzahl der Links auf die darin aufgeführten Sites. Das intellektuelle Qualitätsurteil der Katalogredakteure geht also nicht konform mit der Linkanalyse. Die Anzahl der eingehenden Verbindungen reicht für die Bewertung von Web-Angeboten nicht aus.

Eine mögliche Erweiterung dieser Analyse liegt in der Verfeinerung der Methode. So könnten etwa die Schnittmenge zwischen Verzeichnisdiensten in Bezug auf die referenzierten Seiten gemessen werden. Darauf aufbauend ließe sich untersuchen, ob die gleichen Web-Seiten auch weitgehend in gleiche thematische Kategorien eingeordnet werden.

Eine mögliche Erweiterung liegt in der Einbeziehung internationaler Verzeichnisdienste. Möglicherweise zeigen sich dabei kulturspezifische Unterschiede in der hierarchischen Organisationsstruktur. Eine Anwendung der hier vorgestellten Methodik kann in der Suche innerhalb von Meta-Verzeichnisdiensten liegen. Durch die Link-Analyse können die Seiten unterschiedlicher Dienste zu einer Thematik auf ihre Autorität untersucht werden und die besten dann präsentiert werden.

7 Literaturverzeichnis

- ADAMIC, Lada; HUBERMANN, Bernardo (2001): The Web's Hidden Order. Communications of the ACM 44(9). S. 55-59.
- BAEZA-YATES, Ricardo; RIBEIRO-NETO, Berthier (Hrsg.)(1999): Modern Information Retrieval. Addison-Wesley.
- BARTEL, Torsten (2002): Verbesserung der Usability von WebSites auf der Basis von Web Styleguides, Usability Testing und Logfile-Analysen. Magisterarbeit Universität Hildesheim.
- BEKAVAC, Bernard (1999): Suche und Orientierung im WWW. Verbesserung bisheriger Verfahren durch Einbindung hypertextspezifischer Informationen. Konstanz: Universitätsverlag [Schriften zur Informationswissenschaft Bd. 37].
- CONSTANTOPOULOS, Panos; SOLVBERG, Ingeborg (Hrsg.): Proc 5th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2001).
- EIBL, Maximilian; MANDL, Thomas (2002): Including User Strategies in the Evaluation of Graphic Design Interfaces for Browsing Documents. In: SKALA, Vaclav (Hrsg.): 10th Intl Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG 2002). Pilsen. S. 163-169.
http://wscg.zcu.cz/wscg2002/Papers_2002/B89.pdf
- FAYYAD, Usama (1997): Editorial. In: Data Mining & Knowledge Discovery 1(1). S. 5-10.

- FAYYAD, Usama; UTHURUSAMY, Ramasamy (1996): Data Mining and Knowledge Discovery in Databases. In: Communications of the ACM 39 (11) S. 24.
- FRÖHLICH, Gerhard (2000): Online Informationsvorenthaltung als Strategem wissenschaftlicher Kommunikation. In: ZIMMERMANN, Harald; SCHRAMM, Volker (Hrsg.): Knowledge Management und Kommunikationssysteme: Workflow Management, Multimedia, Knowledge Transfer. Proc 6th Intl Symposium für Informationswissenschaft. (ISI '98). Prag. S. 535-549.
- FROMMHOLZ, Ingo (2001): Categorizing Web Documents in Hierarchical Catalogues. In: Proc 23rd Colloquium on Information Retrieval Research. Darmstadt.
http://www.darmstadt.gmd.de/~frommholz/frommholz_ecir01.pdf
- GIBSON, David; KLEINBERG, Jon; RAGHAVAN, Prabhakar (1998): Inferring Web Communities from Link Topology. In: Proc. 9th ACM Conf on Hypertext and Hypermedia.
<http://citeseer.nj.nec.com/gibson98inferring.html>
- HAWKING, David (2001): Overview of the TREC-9 Web Track. In: VOORHEES & HARMAN 2001.
- KLEINBERG, Jon (1998): Authoritative Sources in a Hyperlinked Environment. In: Proc 9th ACM-SIAM Symposium on Discrete Algorithms.
- KNORZ, Gerhard; KUHLEN, Rainer (Hrsg.): Informationskompetenz – Basiskompetenz in der Informationsgesellschaft. Proc 7. Int. Symposium für Informationswissenschaft. (ISI 2000). Darmstadt.
- KRUSCHWITZ, Udo (2001): Exploiting Structure for Intelligent Web Search. In: Proc Hawai Intl Conf on System Sciences. IEEE.
- LEMPER, R.; MORAN, S. (2000): The Statistical Approach for Link-Structure Analysis (SALSA) and the TKC Effect. In: Proc 9th Intl WWW Conference.
<http://www9.org/w9cdrom/175/175.html>
- MAHOUI, Malika; CUNNINGHAM, Sally (2001): Search Behavior in a Research-Oriented Digital Library. In: CONSTANTOPOULOS & SOLVBERG 2001 S. 13-24.
- MANDL, Thomas (2001): Tolerantes Information Retrieval: Neuronale Netze zur Erhöhung der Adaptivität und Flexibilität bei der Informationssuche. Konstanz: Universitätsverlag [Schriften zur Informationswissenschaft Bd. 39].
- MARCHIONINI, Gary (1995): Information Seeking in Electronic Environments.
- MEGHABGHAB, George (2002): Discovering authorities and hubs in different topological web graph structures. In: Information Processing and Management 38. S. 111-140.
- MUTSCHKE, Peter (2001): Enhancing Information Retrieval in Federated Bibliographic Data Sources Using Author Network Based Stratagems. In: CONSTANTOPOULOS & SOLVBERG 2001. S. 287-299.
- NIE, Jian-Yun.; SIMARD, Michael; FOSTER, George (2001). Multilingual information retrieval based on parallel texts from the web. In: PETERS, Carol (Hrsg.): Cross-Language Information Retrieval and Evaluation. Workshop of the Cross-Language Information Evaluation Forum (CLEF 2000) Lisbon, Portugal, 21.-22. Sept. 2000. Springer [LNCS 2069] S. 188-201.
- PAGE, Larry; BRIN, Sergey; MOTWANI, R.; WINOGRAD, T. (1998): The PageRank Citation Ranking: Bringing Order to the Web.
<http://citeseer.nj.nec.com/page98pagerank.html>
- PIROLI, Peter; PITKOW, James; RAO, Ramana (1996): Silk from a Sow's Ear: Extracting Usable Structures from the Web. In: Proc Conference on Human Factors and

- Computing Systems.
<http://www.acm.org/pubs/articles/proceedings/238386/p118-pirolli/118-pirolli.html>
- RITTBERGER, Marc (2001): Quality Measuring with respect to electronic information markets and particular online databases. In: KENT, A. (Hrsg.): New York, NY: Marcel Dekker, 31, Chapter: 68, S. 274-295.
http://www.inf-wiss.uni-konstanz.de/People/MR/pubs/elis_rittberger.pdf
- SAVOY, Jacques; RASOLOFO, Yves (2000): Report on the TREC-9 Experiment: Link-based Retrieval and Distributed Collections. In: VOORHEES & HARMAN 2000. S. 579.
- SCHLÖGL, Christian (2000): Informationskompetenz am Beispiel einer szionometrischen Untersuchung zum Informationsmanagement. In: KNORZ & KUHLEN 2000: Informationskompetenz - Basiskompetenz in der Informationsgesellschaft. Proc 7. Int. Symposium für Informationswissenschaft. (ISI 2000). Darmstadt. S. 89-112.
- SPINELLO, Richard (2001): An Ethical Evaluation of Web-Site Linking. In: SPINELLO, Richard; TAVANI, Herman (Hrsg.): Readings in CyberEthics. Sudbury, MA et al.: Jones and Bartlett. S. 295-308.
- SULLIVAN, Terry (1997): Reading Reader Reaction: A Proposal for Inferential Analysis of Web Server Log Files. In: Proc 3rd Conf on Human Factors and the Web.
<http://www.pantos.org/ts/papers/rrr.html>
- VOORHEES, Ellen; HARMAN, Donna (2000)(Hrsg.): The Ninth Text REtrieval Conference (TREC 9). NIST Special Publication 500-249.
http://trec.nist.gov/pubs/trec9/t9_proceedings.html
- VOORHEES, Ellen; HARMAN, Donna (2001)(Hrsg.): The Tenth Text REtrieval Conference (TREC 10). NIST Special Publication 500-250.
http://trec.nist.gov/pubs/trec10/t10_proceedings.html
- WALTHER, Ralf (2001): Web Mining. In: Informatik Spektrum 24(1). S. 16-18.
- WEISS, R.; VELEZ, B.; SHELDON, M.; MANPREMPRE, C.; SZILAGYI, P.; DUDA, A.; GIFFORD D. (1996): HyPursuit: A Hierarchical Network Search Engine that Exploits Content-Link Hypertext Clustering. In : Proc Seventh ACM Conference on Hypertext.



Behandlung semantischer Heterogenität durch Metadatenextraktion und Anfragetransfers

Robert Strötgen

Informationszentrum Sozialwissenschaften
Lennéstr. 30
D-53113 Bonn
stroetgen@bonn.iz-soz.de

Zusammenfassung

Die Sonderfördermaßnahme CARMEN¹ (“Content Analysis, Retrieval and Metadata: Effective Networking”) zielte unter anderem darauf ab, die Erweiterung von Recherchen in bibliographischen Fachdatenbanken ins Internet zu verbessern. Dabei war insbesondere die semantische Heterogenität zu behandeln, die durch unterschiedliche Inhaltserschließung in verschiedenen Datenbeständen auftritt. Dazu wurden verschiedene Ansätze wie Metadatenextraktion aus Internetquellen und Anfragetransfers über Cross-Konkordanzen und statistisch erzeugte Relationen gewählt. Dieser Aufsatz stellt die Konzepte und die Implementierung der Metadatenextraktion und der Anfragetransfers sowie die Evaluation der Auswirkungen auf das Retrievalergebnis vor.

Abstract

The project CARMEN (“Content Analysis, Retrieval and Metadata: Effective Networking”) aimed among other goals at improving the expansion of searches in bibliographic databases into Internet searches. We pursued a set of different approaches to the treatment of semantic heterogeneity (meta-data extraction, query translation using statistic relations and cross-concordances). This paper describes the concepts and implementation of this approaches and the evaluation of the impact for the retrieval result.

¹ Gefördert durch das Bundesministerium für Bildung und Forschung im Rahmen des Programms „Global Info“, FKZ 08SFC08 3.



1 Semantische Heterogenität

In der heutigen dezentralen Informationswelt werden Benutzer mit einer Vielzahl dezentraler Informationssysteme und Datenbestände mit ganz unterschiedlichen Inhaltserschließungsverfahren konfrontiert. In diesem Zusammenhang tritt semantische Heterogenität² auf, wenn unterschiedliche Datenbestände, die über eine Suchfunktion gemeinsam zugänglich gemacht werden, verschiedene Dokumentationssprachen benutzen, wenn Metadaten unterschiedlich oder überhaupt nicht erfasst werden oder wenn intellektuell aufgearbeitete Quellen mit in der Regel vollständig unerschlossenen Internetdokumenten zusammentreffen.

Standardisierungsbestrebungen wie die der Dublin Core Metadata Initiative sind eine wichtige Voraussetzung für eine Verbesserung der Verbindung von Datenbeständen, aber sie erfordern ein hierarchisches Modell der Kooperation, das von allen Beteiligten akzeptiert wird. Wegen der verschiedenen Interessen der unterschiedlichen Partner lässt sich ein solches Modell kaum verwirklichen. Vielmehr ist davon auszugehen, dass durch „anarchische Tendenzen“ die Unterschiede bei der Erstellung, Erschließung und Verbreitung von Dokumenten eher zunehmen werden. [Krause '96, Krause/Marx '00]

Im Projekt CARMEN wurde dieses Problem einerseits durch die automatische Extraktion von Metadaten aus Internetdokumenten und andererseits durch Systeme zur Transformation von Anfragen über Cross-Konkordanzen und statistisch erzeugte Relationen angegangen.

Zunächst wird in diesem Aufsatz die Extraktion von Meta-Daten aus Internetquellen beschrieben (Kap. 2). Daran anschließend wird das Vorgehen bei der Erstellung von Cross-Konkordanzen und statistischen Relationen dargestellt (Kap. 3). Ähnliche Verfahren wurden in den Projekten AIR/PHYS [Biebricher et al. '88] verwendet und werden auch für die „EuroSpider“³-Systeme für multilinguales Retrieval genutzt. [Braschler/Schäuble '98, '00]

Die Verfahren zum Transfer von Anfragen (Kap. 4) berücksichtigen, dass eine Veränderung der Daten in den abzufragenden Datenbanken oft nicht möglich ist. Anfrageerweiterung wurde im Zusammenhang mit „relevance feedback“ diskutiert. [Harman '88] Hier dient es der Übersetzung zwischen De-

² Der Begriff semantische Heterogenität ist hier anders zu verstehen als in der Diskussion um die technischen Probleme der Behandlung verschiedener DBMS mit unterschiedlichen Schemata. [Bright et al. '94, Hull '97].

³ <http://www.eurospider.ch/>.

skriptoren oder Notationen verschiedenen Dokumentationssprachen während des Retrievals. In Retrievaltests (Kap. 5) wurde die Auswirkung der Transferverfahren auf das Rechercheergebnis untersucht.

2 Extraktion von Meta-Daten

2.1 Ansatz

Das Ziel bei der Extraktion von Meta-Daten ist die Anreicherung schlecht oder gar nicht inhaltlich erschlossener Quellen mit wichtigen Meta-Daten, wie z.B. Autor, Titel, Schlagwörter oder Abstract, schon während des Einsammelns von Dokumenten – während der unten beschriebene Transfer von Anfragen zum Suchzeitpunkt durchgeführt wird. Diese generierten Meta-Daten werden den Dokumenten hinzugefügt und stehen dann für das Retrieval zusammen mit sicheren Meta-Daten zur Verfügung, sie haben allerdings ein niedrigeres Gewicht.

Die entwickelten Heuristiken für die Extraktion von Meta-Daten hängen stark von Dateiformaten, vom Fachgebiet, von Eigenschaften einer Website und vom Layout der Seiten ab. Ein zuverlässiger und fach- wie institutionenunabhängiger Ansatz ist bisher nicht verfügbar. So lange sich die Konventionen zum Erstellen von HTML-Dokumenten so schnell ändern und ein „semantisches Web“ nicht vorliegt, können nur beschränkte und temporär gültige Lösungen entwickelt werden.

Relevante Internet-Dokumente für die Mathematik liegen meist im PostScript-Format vor. Die hier erschlossenen Dokumente (Preprints und Dissertationen) werden also in einem unstrukturierten Dateiformat gespeichert, enthalten allerdings hochwertige Meta-Daten, die durch Schlüsselwörter und Layout-Informationen erkannt werden können. Durch die Anwendung von Werkzeugen wie „Prescript“⁴ der New Zealand Digital Library wurde es möglich, die PostScript-Dokumente zu transformieren und zu analysieren. Dabei wurden der Abstract, Schlagwörter und Klassifikationen erfolgreich extrahiert.

Internet-Dokumente aus den Sozialwissenschaften liegen in aller Regel als HTML-Dokumente vor, also in einem strukturierten Dateiformat. Allerdings werden die Möglichkeiten, mit HTML Inhalte auszuzeichnen, oft nur für die Formatierung genutzt. Meta-Tags werden kaum genutzt, und die Dokumente sind syntaktisch oft fehlerhaft. Verschiedene Institutionen benutzen unter-

⁴ <http://www.nzdl.org/html/prescript.html>.

schiedliche Stile bei der Erstellung ihrer Dokumente, und viele Dokumente enthalten nicht einmal Informationen über den Autor oder die Institution selbst. Dadurch wird die Extraktion von Meta-Daten sehr erschwert.

Da die Analyse der (häufig fehlerhaften) HTML-Dokumente sehr fehleranfällig ist, wurden die Dokumente in XHTML konvertiert und dabei syntaktisch bereinigt. Dadurch wurde es möglich, bei der Extraktion bereits vorhandene XML-Werkzeuge zu nutzen und die Heuristiken mit der Anfragesprache XPath zu erstellen. [Strötgen/Kokkelink '01].

Das folgende Beispiel stellt den Algorithmus für die Extraktion und die Gewichtung des Titels dar (<x> gibt dabei eine interne Methoden-Nummer an, [x] das Gewicht der Meta-Daten). Zugrunde liegt dabei eine vorher durchgeführte Vorstrukturierung der Dokumente mit XPath.

```
If (<title> vorhanden && <title> enthält nicht „untitled“ && HMAX
    vorhanden){
    /* 'enthält nicht „untitled“' wird case insensitive im <titel>
        als Substring gesucht */
    If (<title>==HMAX) {
        <1> Titel[1,0]=<title>
    } elseif (<title> enthält HMAX) {
        /* 'enthält' meint hier immer case insensitive als Substring */
        <2> Titel[0,8]=<title>
    } elseif (HMAX enthält <title>) {
        <3> Titel[0,8]=HMAX
    } else {
        <4> Titel[0,8]=<title> + HMAX
    }
} elseif (<title> vorhanden && <title> enthält nicht „untitled“ &&
    S vorhanden) {
    /* d.h. <title> vorhanden UND es existiert ein Eintrag S mit
        //p/b, //i/p usw. */
    <5> Titel[0,5]=<title> + S
} elseif (<title> vorhanden) {
    <6> Titel[0,5]=<title>
} elseif (<Hx> vorhanden) {
    <7> Titel[0,3]=HMAX
} elseif (S vorhanden)
{
    <8> Titel[0,1]= S
}
}
```

Abbildung 1: Heuristik für die Extraktion von Titeln

2.2 Evaluation

Für die Sozialwissenschaften wurde ein Testkorpus mit 3661 HTML-Dokumenten von den Web-Servern verschiedener relevanter Institutionen gesammelt. Von diesen Dokumenten enthielten 96% einen korrekt ausgezeich-

neten Titel; 17,7% der übrigen Dokumente enthielt einen fehlerhaft codierten Titel, die übrigen überhaupt keinen. Nur 25,5% enthalten Schlagworte, die alle korrekt ausgezeichnet wurden. Nur 21% enthalten einen richtig ausgezeichneten Abstract, 39,4% der übrigen Dokumente enthalten einen Abstract, der auf andere Weise im Dokument eingebaut wurde. Diese Voruntersuchung ist die Grundlage der Evaluation – denn da keine Methoden zur automatischen Inhaltserschließung angewendet werden sollten, können Meta-Daten nur da extrahiert werden, wo sie auch im Dokument vorhanden sind.

Für die Evaluation wurde eine repräsentative Stichprobe im Umfang von 360 Dokumenten gebildet, für die intellektuell die Relevanz der extrahierten Meta-Daten bestimmt wurde. Dabei wurden vier Relevanzstufen genutzt: hohe Präzision und Vollständigkeit, hohe Präzision für einen Teil der Extraktion, aber unvollständig und/oder auch Extraktion nicht relevanter Anteile; nicht relevant; nicht bewertbar.

Von den extrahierten Titeln sind 80% von mittlerer oder hoher Qualität, fast 100% der gefundenen Schlagworte sind von hoher Qualität und etwa 90% der extrahierten Abstracts sind von hoher oder mittlerer Qualität. [Binder et al. '02]

3 Semantische Relationen

Semantische Relationen zwischen Elementen von Dokumentationssprachen oder auch von solchen Elementen zu Freitexttermen wurden im Projekt CARMEN intellektuell erzeugt („Cross-Konkordanzen“) und über statistische Verfahren generiert.

3.1 Cross-Konkordanzen

Für die Bereiche Mathematik, Physik und Sozialwissenschaften wurden von Fachexperten intellektuell semantische Relationen zwischen verschiedenen Thesauri und Klassifikationen erstellt. Diese Verknüpfungen verbinden jeweils zwei Dokumentationssprachen miteinander, die Relationen wurden als Äquivalenz, Ober-/Unterbegriffsrelation und Ähnlichkeitsrelation erfasst und mit Gewichten (hoch, mittel, niedrig) versehen.

Für die Erstellung derartiger Cross-Konkordanzen wurden zwei Werkzeuge eingesetzt: das Web-basierte CarmenX⁵ für die Konkordanzen zwischen Klas-

⁵ <http://www.bibliothek.uni-regensburg.de/projects/carmen12/>.

sifikationen und das als semantisches Netzwerk organisierte SIS/TMS⁶ für die Relationen zwischen Thesauri.

Einmal erarbeitete Konkordanzen ermöglichen sichere Übergänge zwischen verschiedenen Erschließungssystemen, ihre Erstellung und Wartung ist aber mit hohem finanziellen und zeitlichen Aufwand verbunden. Außerdem sind viele Dokumente, vor allem im Internet, überhaupt nicht mit einem kontrollierten Vokabular erschlossen. Daher sind ergänzend oder alternativ zusätzliche automatische Verfahren zu nutzen.

3.2 Statistisch erzeugte Relationen

Statistische Methoden ermöglichen die Erzeugung von semantischen Relationen auf der Grundlage vorhandener Dokumentbestände. Im beschriebenen Projekt wurde die Analyse von Wort-Kookkurrenzen gewählt. In den vergangenen 20 Jahren wurden vor allem im Kontext der automatischen Inhaltsererschließung verschiedene Verfahren zur Kookkurrenzanalyse erprobt, [Biebricher et al. '88, Ferber '97] wobei die bedingte Wahrscheinlichkeit und der Äquivalenzindex mit die besten Ergebnisse lieferten und daher auch hier Anwendung fanden. Voraussetzung ist ein Parallelkorpus, der zwei Dokumentationssprachen verbindet [Mandl '99].

Für die Erzeugung semantischer Relationen, die für einen Transfer zwischen Fachdatenbanken und Internetquellen genutzt werden können, lässt sich das benötigte Parallelkorpus nur mit Mühe bereitstellen. Internetquellen, insbesondere aus den Sozialwissenschaften, sind in aller Regel überhaupt nicht inhaltlich erschlossen, schon gar nicht mit einem kontrollierten Vokabular. Da sich erst recht kaum Internetdokumente finden lassen, die mit mehreren Dokumentationssprachen erschlossen sind, ließ sich über Internetquellen kein geeignetes Parallelkorpus direkt erstellen.

Ein weiteres Ziel war die Verbindung von Dokumentationssprachen mit Freitexttermen, die im Gegensatz zu einem Thesaurus oder einer Klassifikation nicht kontrolliert sind und daher einer Eingrenzung und Vorbehandlung bedürfen.

Aus dem Bereich Sozialwissenschaften wurde ein Testkorpus mit etwa 6000 HTML-Dokumenten von wissenschaftlichen Einrichtungen im Internet zusammengestellt. Über diesen Bestand wurde ein Parallelkorpus simuliert, in-

⁶ <http://www.ics.forth.gr/proj/isst/Systems/sis-tms.html>.

dem den Dokumenten einerseits über einfache Verfahren der automatischen Inhaltserschließung Deskriptoren aus dem Thesaurus Sozialwissenschaften und andererseits über den Volltextindexierer Fulcrum SearchServer Freitextterme zugeordnet wurden. Die Freitextterme wurden mit einem Porter-Stemmer⁷ vorbehandelt und über Schwellenwerte wurde die Zahl der zugeordneten Freitextterme für ein Dokument begrenzt. Auf diese Weise wurde ein Korpus bereitgestellt, dessen Dokumente durch zwei verschiedene (Dokumentations-)Sprachen erschlossen waren und der nun wie ein Parallelkorpus behandelt wurde (siehe Abb. 2).

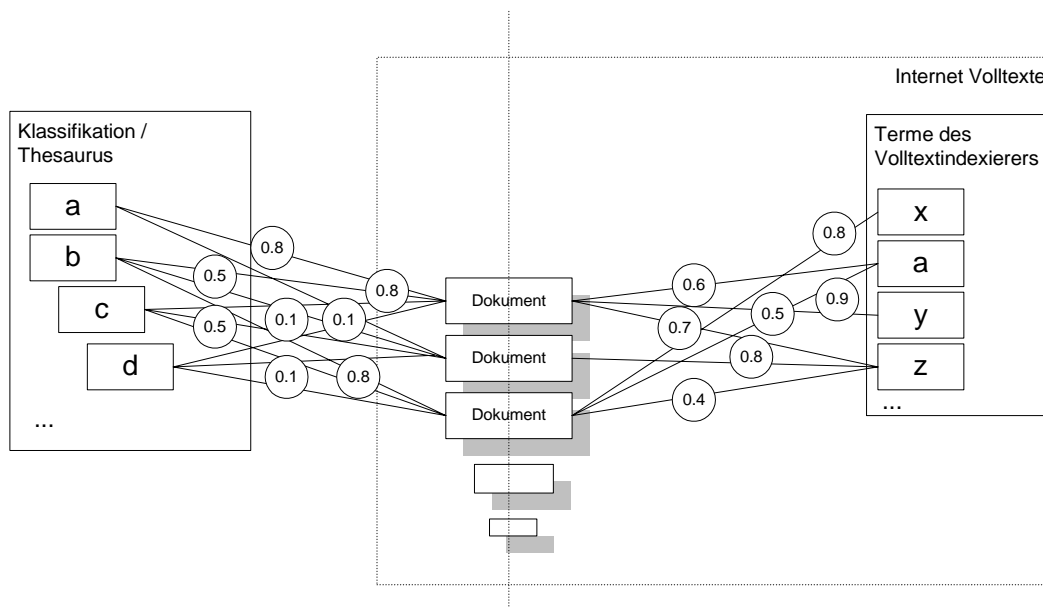


Abbildung 2: Parallelkorpus-Simulation mit vagen Deskriptoren und Volltexttermen

Für die Realisierung statistischer Transfermodule wurde auf das Werkzeug „Jester“ (**J**ava **E**nvironment for **S**tatistical **T**ransf**E**Rs) zugegriffen, das im Kontext des Projekts Elvira II erstellt wurde. Jester führt die Berechnung der statistischen Transferbeziehungen durch und unterstützt dabei den Benutzer bei der Auswahl nötiger Parameter.

Als Maß für den statistischen Zusammenhang zweier Terme wird die „bedingte Wahrscheinlichkeit“, mit der zwei Terme gemeinsam auftreten, berechnet. Für einen Deskriptor **a** bestimmt $P(a)$ die Wahrscheinlichkeit, dass **a** einem Dokument zugeordnet ist. Sie lässt sich ermitteln, indem die Zahl der Dokumente, in denen **a** auftritt durch die Gesamtzahl der Dokumente geteilt wird. $P(a \wedge b)$ bezeichnet die Wahrscheinlichkeit, dass zwei Terme **a** und **b** gemeinsam auftreten. Ist das Auftreten der beiden Terme unabhängig vonein-

⁷ Vgl. <http://www.tartarus.org/~martin/PorterStemmer/>.

ander, so sollte in etwa $P(a \wedge b) = P(a) * P(b)$ gelten. Weichen diese Werte erheblich voneinander ab, so liegt ein systematischer Zusammenhang zwischen den Termen vor. Treten sie häufiger gemeinsam auf, als vorhergesagt, so sind sie wahrscheinlich nah verwandt. Treten sie deutlich seltener gemeinsam auf, so liegt nahe, dass sie eine gegensätzliche Beziehung haben und einander ausschließen.

Anstatt nur diese symmetrische Beziehung zu betrachten, die bei geeigneter Wahl eines Schwellenwertes die Ermittlung von Termpaaren erlaubt, die quasi synonym verwendet werden, kann auch die gerichtete bedingte Wahrscheinlichkeit betrachtet werden. Hier wird nur betrachtet, wie wahrscheinlich es ist, dass, wenn Deskriptor **a** vergeben wurde, ebenfalls Deskriptor **b** auftritt. $P(b) / P(a \wedge b)$ beschreibt diesen Zusammenhang. Je größer dieser Wert ist, umso höher ist die gerichtete Abhängigkeit, die z.B. auftritt, falls **a** semantisch ein Unterbegriff von **b** ist.

Jester führt den Bearbeiter schrittweise durch die notwendigen Abläufe und unterstützt bei der Auswahl von Schwellenwerten. Auf diese Weise kann direkt bei der Manipulation der Parameter die Auswirkung beobachtet werden und interaktiv eine optimale Auswahl getroffen werden. Darüber hinaus stellt es eine Reihe von Heuristiken bereit, die bei der Behandlung systematischer Fehler wie Ausreißer angewendet werden können. [Hellweg '02]

Das Ergebnis der auf diesen Parallelkorpora durchgeführten Wort-Kookkurrenz-Analysen sind Term-Term-Matrizen, in der die gefundenen semantischen Relationen zwischen Deskriptoren und Freitexttermen für die spätere Nutzung beim Transfer von Anfragen bereitgehalten werden. [Hellweg et al. '01]

4 Transfer von Anfragen

Die semantischen Relationen waren nun für ein Suchsystem nutzbar zu machen. Da die genutzten verteilten Datenbanken nicht verändert werden sollten, sind die Relationen zur Manipulation der Suchanfragen zwischen Benutzerschnittstelle und Datenbank anzusiedeln. Die Suchanfrage soll also nicht - wie sonst in verteilten Sucharchitekturen, z.B. Meta-Suchmaschinen im WWW verbreitet - unverändert an die einzelnen Datenbanken weitergeleitet werden. Die Anfrage soll auch nicht einmalig erweitert und dann an alle Datenbanken gleich gestellt werden.

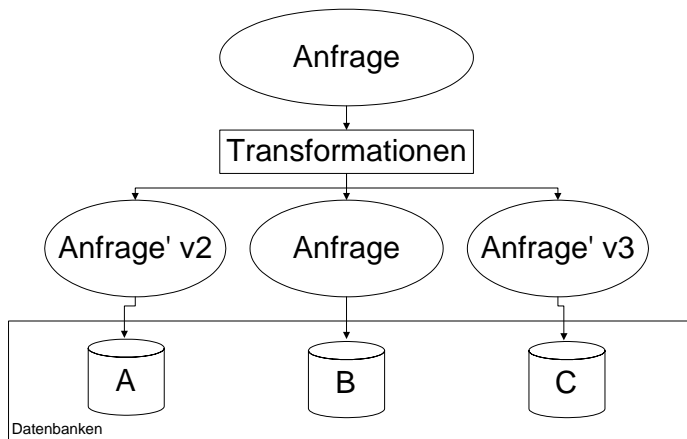


Abbildung 3: „Zwei-Schritt-Verfahren“

Stattdessen wird das im Informationszentrum Sozialwissenschaften entwickelte „Zwei-Schritt-Verfahren“⁸ angewendet (siehe Abb. 3). Dabei werden datenbankspezifisch Anfragen generiert, die die jeweilige Inhaltserschließung der Datenbank berücksichtigen. Ist die Anfrage mit Hilfe einer Dokumentationssprache formuliert, die in einer der beteiligten Datenbanken genutzt wird, so wird die Anfrage unverändert an diese Datenbank gestellt. Für eine Datenbank, die eine andere Dokumentationssprache nutzt, wird die Anfrage über semantische Relationen zwischen beiden Dokumentationssprachen übersetzt.

Für das Projekt CARMEN waren die in Java entwickelten Softwaremodule, die diese Übersetzung der Anfrage leisten, in die Gesamtarchitektur einzubauen. Das an der Universität Dortmund entwickelte Retrievalsystem HyRex [Fuhr et al. '00] wurde dafür erweitert. Teilanfragen, für die ein Transfer angewendet werden kann (also z.B. nach Schlagwort) werden aus der in der Anfragesprache XIRQL vorliegenden Gesamtanfrage herausgetrennt. Die XIRQL-Teilanfrage wird in XML codiert und per Http-Request an den Servlet-Server gestellt, dort übersetzt und als Http-Response zurückgeschickt (siehe Abb. 3). In HyRex wird schließlich die veränderte Teilanfrage wieder in die Gesamtanfrage eingebaut. Sind mehrere Datenbanken beteiligt, ist dieser Schritt ggf. für jede Datenbank zu wiederholen.

⁸ Angewendet in den Projekten ELVIRA, CARMEN, ViBSoz und ETB, vgl. <http://www.gesis.org/Forschung/Informationstechnologie/Heterogenitaet.htm>.

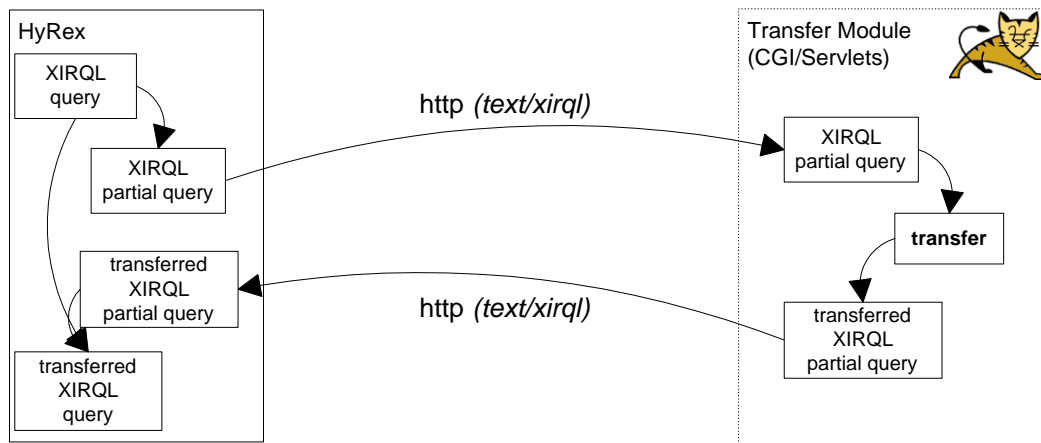


Abbildung 3: Anfrage-Transfer-Architektur

5 Evaluation

Für die Evaluation der Anfragentransfers über statistische Relationen wurden etwa 10000 HTML-Dokumente aus den Sozialwissenschaften mit dem Volltextindexierer Fulcrum indexiert (HyRex stand leider nicht rechtzeitig zum Test zur Verfügung). Die Ausnutzung der Gewichtsinformationen, die in HyRex genutzt werden, war dabei nicht möglich.

Das Test-Szenario geht von einer Suche in der Literaturlatenbank SOLIS⁹ aus, die mit dem Thesaurus Sozialwissenschaften erschlossen ist. Eine Anfrage mit Deskriptoren aus diesem Thesaurus soll dann in einer Internet-Anfrage mit Freitexttermen erweitert werden. Die Thesaurus-Deskriptoren werden dabei also um Freitextterme erweitert.

Für den Test wurden zu drei Bereichen aus den Sozialwissenschaften (Frauenforschung, Migration und Industriesoziologie) jeweils zwei Anfragen gestellt. Dabei wurde zunächst mit der unveränderten SOLIS-Anfrage in Internet-Dokumenten gesucht und anschließend die transferierte Anfrage gestellt. Zwei Beispiele sollen hier kurz vorgestellt werden:

Eine Anfrage nutzte den Deskriptor „Dominanz“ und lieferte 16 relevante Dokumente. Die transferierte Anfrage enthielt neu zusätzliche Terme: „Messen“, „Mongolei“, „Nichtregierungsorganisation“, „Flugzeug“, „Datenaustausch“, „Kommunikationsraum“, „Kommunikationstechnologie“, „Medienpädagogik“, „Wüste“. Die übersetzte Anfrage lieferte 14 zusätzliche Dokumente, von denen 7 relevant waren (50%, Zugewinn 44%). In diesem Beispiel

⁹ <http://www.gesis.org/Information/SOLIS/>

konnten mit wenig Ballast zusätzliche relevante Dokumente gefunden werden.

Die genutzten Relationen sind mit großer Vorsicht zu interpretieren. „Wüste“ und „Mongolei“ werden nicht ergänzt, weil ein besonderes Dominanzproblem in der mongolischen Wüste angenommen wird. Es gibt lediglich ein Dokument, in dem eine Exkursion von Frauenrechtlerinnen nach China mit einem Zwischenstopp in der Mongolei beschrieben wird und dieses Dokument durch ungünstige Gewichtung einen hohen Einfluss auf die Relationen erhalten hat. Es gibt aber durchaus andere Fälle, in denen semantische Relationen durchaus Hinweise auf Problemfelder geben, z.B. bei statistischen Relationen zwischen mathematischen und physikalischen Klassifikationen, die mathematische Methoden mit physikalischen Anwendungsbereichen verbinden.

Ein anderes, weniger erfolgreiches Beispiel ist die Suche nach dem Deskriptor „Leiharbeit“. Zunächst wurden 10 relevante Dokumente gefunden. In der transferierten Anfrage wurden drei zusätzliche Terme hinzugefügt: „Arbeitsphysiologie“, „Organisationsmodell“, „Risikoabschätzung“. Von den 10 zusätzlich gefundenen Dokumenten waren nur 2 relevant (20%, Zugewinn 20%). In diesem Beispiel wurden mit erheblichem Ballast (80%) kaum zusätzliche Dokumente gefunden.

Zusammenfassend kann festgehalten werden, dass in allen transferierten Anfragen zusätzliche relevante Dokumente gefunden wurden. Die Precision der zusätzlichen Treffer liegt zwischen 13% und 55%. Ohne systematische Zusammenhänge bereits erkannt zu haben, wurden eher erfolgreiche und eher schwache Auswirkungen der Anfragetransfers vorgefunden.

6 Zusammenfassung und Ausblick

Es hat sich gezeigt, dass prinzipiell Meta-Daten extrahiert werden können. Die Extraktions-Werkzeuge wurden in den Gatherer des CARMEN-Projects „CARA“ eingebaut und stehen auch für andere Anwendungen zur Verfügung. Wegen der Kurzlebigkeit und des hohen Wartungsbedarfs der entwickelten Heuristiken scheint es aber fraglich, ob dieser hohe Aufwand aufgebracht werden kann.

Der Transfer von Anfragen unter Ausnutzung statistisch erzeugter Relationen hat sich grundsätzlich als brauchbar erwiesen, um die Ergebnisse der Suche zu verbessern. Allerdings bleiben einige Punkte offen. Zu klären ist beispielsweise, wie die Korpora und die Verfahren verbessert werden müssen,

um zu besseren Term-Term-Matrizen zu kommen. Außerdem wären die Anfragetransfers mit intellektuell erstellten Cross-Konkordanzen zum Vergleich heranzuziehen. Schließlich ist in echten Benutzertests zu evaluieren, welche Auswirkungen die Transfermodule im interaktiven Retrieval haben, wie sie von Benutzern sinnvoll parametrisiert werden können und welche Probleme und Irritationen beim Benutzer auftreten können.

Die im Projekt CARMEN erzeugten Softwaremodule und Term-Term-Matrizen werden interessierten Anwendern zur Verfügung gestellt. Im IZ Sozialwissenschaften finden sie in anderen Diensten wie ViBSoz und ETB Verwendung, weitere Dienste wie der Informationsverbund Bildung - Sozialwissenschaften - Psychologie werden folgen.

7 Literaturverzeichnis

- [Biebricher et al. '88] Biebricher, P.; Fuhr, N.; Lustig, G.; Schwantner, M.; Knorz, G.: The Automatic Indexing System AIR/PHYS. From Research to Application. In: Chiaramella, Y. (Hrsg.): SIGIR'88, Proceedings of the 11th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Grenoble, France, June 13-15, 1988. ACM, 1988. S. 333-342.
- [Binder et al. '02] Binder, G.; Marx, J.; Mutschke, P.; Strötgen, R.; Plümer, J.; Kokkelink, S.: Heterogenitätsbehandlung bei textueller Information verschiedener Datentypen und Inhaltserschließungsverfahren. (IZ-Arbeitsbericht; Nr. 24) Bonn: IZ Sozialwissenschaften, 2002.
- [Braschler/Schäuble '98] Braschler, M.; Schäuble, P.: Multilingual Information Retrieval Based on Document Alignment Techniques. In: Nikolaou, C.; Stephanidis, C. (Hrsg.): Research and Advanced Technology for Digital Libraries, Second European Conference, ECDL '98, Heraklion, Crete, Greece, September 21-23, 1998, Proceedings. (Lecture Notes in Computer Science): Springer, 1998. S. 183-197.
- [Braschler/Schäuble '00] Braschler, M.; Schäuble, P.: Using Corpus-Based Approaches in a System for Multilingual Information Retrieval. In: Information Retrieval 2000, 3, Nr. 3; S. 273-284.
- [Bright et al. '94] Bright, M.W.; Hurson, A.R.; Pakzad, S.H.: Automated Resolution of Semantic Heterogeneity in Multidatabases. In: ACM Transactions on Database Systems (TODS) 1994, 19, Nr. 2; S. 212-253.
- [Ferber '97] Ferber, R.: Automated Indexing with Thesaurus Descriptors: A Co-occurrence Based Approach to Multilingual Retrieval. In: Peters, C.; Thanos, C. (Hrsg.): Research and Advanced Technology for Digital Libraries. First European Conference, ECDL '97, Pisa, Italy, 1-3 September, Proceedings. (Lecture Notes in Computer Science): Springer, 1997. S. 233-252.
- [Fuhr et al. '00] Fuhr, N.; Großjohann, K.; Kokkelink, S.: CAP7: Searching and Browsing in Distributed Document Collections. In: Borbinha, J.L.; Baker, T. (Hrsg.): Research and Advanced Technology for Digital Libraries, 4th European Conference, ECDL 2000,

- Lisbon, Portugal, September 18-20, 2000, Proceedings. (Lecture Notes in Computer Science): Springer, 2000. S. 364-367.
- [Harman '88] Harman, D.: Towards Interactive Query Expansion. In: Chiaramella, Y. (Hrsg.): SIGIR'88, Proceedings of the 11th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Grenoble, France, June 13-15, 1988. ACM, 1988. S. 321-331.
- [Hellweg '02] Hellweg, H.: Einsatz von statistisch erstellten Transferbeziehungen zur Anfrage-Transformation in ELVIRA. In: Krause, J.; Stempfhuber, M. (Hrsg.): Integriertes Retrieval in heterogenen Daten. Text-Fakten-Integration am Beispiel des Verbandinformationssystems ELVIRA. (Forschungsberichte des IZ Sozialwissenschaften) Bonn: IZ Sozialwissenschaften, 2002.
- [Hellweg et al. '01] Hellweg, H.; Krause, J.; Mandl, T.; Marx, J.; Müller, M.N.O.; Mutschke, P.; Strötgen, R.: Treatment of Semantic Heterogeneity in Information Retrieval. (IZ-Arbeitsbericht; Nr. 23) Bonn: IZ Sozialwissenschaften, 2001.
- [Hull '97] Hull, R.: Managing Semantic Heterogeneity in Databases. A Theoretical Perspective. In: ACM Symposium on Principles of Databases. Proceedings. ACM, 1997. S. 51-61.
- [Krause '96] Krause, J.: Informationserschließung und -bereitstellung zwischen Deregulation, Kommerzialisierung und weltweiter Vernetzung ("Schalenmodell"). (IZ-Arbeitsbericht; Nr. 6) Bonn: IZ Sozialwissenschaften, 1996.
- [Krause/Marx '00] Krause, J.; Marx, J.: Vocabulary Switching and Automatic Metadata Extraction or How to Get Useful Information from a Digital Library. In: Information Seeking, Searching and Querying in Digital Libraries: Pre-Proceedings of the First DELOS Network of Excellence Workshop. Zürich, Switzerland, December, 11-12, 2000. Zürich, 2000. S. 133-134.
- [Mandl '99] Mandl, T.: Effiziente Implementierung von statistischen Assoziationen im Text-Retrieval. In: Ockenfeld, M.; Mantwill, G.J. (Hrsg.): Information und Region; 51. Jahrestagung der Deutschen Gesellschaft für Informationswissenschaft und Informationspraxis e.V., DGI, Hamburg, 21. bis 23. September 1999; Proceedings. (Tagungen der Deutschen Gesellschaft für Informationswissenschaft und Informationspraxis; 1) Frankfurt am Main: DGI, 1999. S. 159-172.
- [Strötgen/Kokkelink '01] Strötgen, R.; Kokkelink, S.: Metadatenextraktion aus Internetquellen: Heterogenitätsbehandlung im Projekt CARMEN. In: Schmidt, R. (Hrsg.): Information Research & Content Management: Orientierung, Ordnung und Organisation im Wissensmarkt; 23. Online-Tagung der DGI und 53. Jahrestagung der Deutschen Gesellschaft für Informationswissenschaft und Informationspraxis e.V., DGI, Frankfurt am Main, 8. bis 10. Mai 2001; Proceedings. (Tagungen der Deutschen Gesellschaft für Informationswissenschaft und Informationspraxis; 4) Frankfurt am Main: DGI, 2001. S. 56-66.



Wider einige gängige Ansichten zur juristischen Informationserschließung

Bettina Mielke

Landgericht Regensburg
93066 Regensburg
bettina.mielke@lg-r.bayern.de

Zusammenfassung

Ausgehend von einer Betrachtung in der Rechtsinformatik geläufiger Annahmen zur juristischen Informationserschließung beschreibt der Beitrag wesentliche Ergebnisse einer empirischen Studie der Retrievaleffektivität von Recherchen in juristischen Datenbanken. Dabei steht die Frage nach der Notwendigkeit einer intellektuellen Erschließung einerseits, der Effektivität der sogenannten Stichwortsuche andererseits im Mittelpunkt. Die Ergebnisse der Studie, bei der auch ein Vergleich zwischen einem Informationssystem auf der Basis eines Booleschen Retrievalmodells mit einem System auf der Basis statistischer Verfahren vorgenommen wurde, legen den Schluss nahe, dass in der rechtsinformatischen Fachliteratur analytisch begründete Annahmen wie die Gefahr zu großer Antwortmengen bei der Stichwortsuche empirisch nicht zu belegen sind. Auch zeigt sich keine Überlegenheit intellektueller Erschließungsverfahren (Beschlagwortung) gegenüber der automatischen Indexierung, im Gegenteil führt der Einsatz eines statistischen Verfahrens bei identischer Dokumentkollektion zu einer höheren Wiedergewinnungsrate (recall).

Abstract

Starting from a short analysis of wide-spread assumptions about law information systems this paper describes the main results of an evaluation study concerning the effectiveness of searches in juridical databases. The need for intellectual indexing as well as the effectiveness of keyword-based search strategies are at the heart of this study. Its results which also comprise a comparison of a retrieval system based on Boolean logic with a statistical search engine show that analytical assumptions about too large answer sets for keyword searches cannot be verified empirically. Furthermore, no indication for advantages of intellectual indexing can be seen. To the contrary, using a sta-



tistical search engine results in higher recall values search the same document collection.

1 Einleitung

Internet- bzw. webbasierte Informationssysteme sind heute auch im Bereich der juristischen Fachinformation in vielfältiger Weise verfügbar. Zwar haben sie die juristische Praxis noch nicht in einer mit anderen Fachdisziplinen vergleichbaren Weise durchdrungen, werden aber mehr und mehr zum alltäglichen Arbeitswerkzeug des praktizierenden Juristen.

Dabei darf nicht darüber hinweggesehen werden, dass der Aspekt der *generellen Verfügbarkeit* durch Zugang zum WWW und damit auch zu juristischen Informationssystemen von jedem Arbeitsplatz aus sowie der Aspekt der Verbesserung der hierfür entwickelten Benutzerschnittstellen im WWW bisher wenig an der grundsätzlichen Problematik der Erschließung relevanter Fachinformation geändert haben dürften.

Der folgende Beitrag befasst sich mit einigen gängigen Thesen zur juristischen Informationserschließung und stellt diesen die Ergebnisse einer empirischen Evaluierungsstudie gegenüber. Dabei zeigt sich, dass sich die allein aufgrund von analytischen Betrachtungen gewonnenen Einschätzungen empirisch nicht bestätigen lassen.

2 Literatursuche als notwendiger Bestandteil der juristischen Praxis

Die Literatursuche ist ein notwendiger Bestandteil der juristischen Praxis. Sie beginnt mit der Suche nach den einschlägigen und zu einem bestimmten Zeitpunkt gültigen Gesetzesbestimmungen, was bei selten vorkommenden Nebengesetzen durchaus aufwändig sein kann. Auch die Suche nach Gerichtsentscheidungen und Literaturbeiträgen ist in nahezu allen juristischen Tätigkeitsfeldern unumgänglich - man denke allein an die Haftung des Rechtsanwalts, der seinen Mandanten falsch berät, da er die neueste obergerichtliche Rechtsprechung nicht kennt.

Die traditionell wichtigste juristische Sekundärliteratur in Deutschland stellen Kommentare und Handbücher dar, die Gerichtsentscheidungen und wissenschaftliche Rechtsliteratur auswerten und einen ersten Überblick zu den vor-

handenen Rechtstexten geben.¹ Daneben besteht ein überaus vielfältiges Spektrum von Fachzeitschriften und sonstigen Sammlungen, in denen Gerichtsentscheidungen und Aufsätze veröffentlicht sind und die über verschiedene Register inhaltlich erschlossen sind, z.B. die Fundhefte für Zivilrecht, Systematischer Nachweis der Deutschen Rechtsprechung und Zeitschriftenaufsätze oder die Karlsruher Juristische Bibliographie.²

Juristische Informationssysteme stellen nicht nur eine Alternative zur traditionellen Informationsbeschaffung dar, sondern eröffnen neue, direkt vom Arbeitsplatz aus erreichbare Informationsquellen, die hinsichtlich ihrer Leistungsfähigkeit über die traditionellen Sammlungen und Nachschlagewerke weit hinausgehen.

3 Einige gängige Thesen zur Informationserschließung

3.1 Notwendigkeit einer intellektuellen Erschließung

Das elektronische Publizieren von rechtlichen Texten, also Normen, Judikatur und Literatur, ist nach wie vor eines der wichtigsten Themen der Rechtsinformatik.³ Dabei geht die deutschsprachige Rechtsinformatik überwiegend davon aus, dass die intellektuelle Erschließung einen Mehrwert bringt und die automatische Erschließung nicht „die sehr hohe Recall und Präzision der intellektuellen Erschließung“⁴ erziele. So zeige z.B. die „hohe Akzeptanz einer juristischen Erschließung in Österreich“ deutlich, „dass außerhalb des öffentlichen Sektors (und oft auch dort) ohne einen Index juristische Arbeit nur erschwert möglich ist. Die nächsten Jahre werden eine zunehmende Unterstützung und Verbesserung dieser Indexierung durch neue Methoden der Rechtsinformatik bringen.“⁵ Diese Behauptungen werden dabei ohne empirische Grundlage getroffen, da Evaluierungsstudien zu diesem Thema fehlen. Im deutschsprachigen Raum gibt es bislang nur Endnutzerbefragungen, die nach

¹ Vgl. etwa Möllers, JuS 2000, S. 1203 (1208).

² Vgl. Mielke 2000, S. 15 f.; Möllers, JuS 2000, S. 1203 (1205 ff.).

³ Elektronisches Publizieren und Rechtsdatenbanken, Vorbemerkungen der Herausgeber, in: Schweighofer/Menzel/Kreuzbauer 2001, S. 97 ff.

⁴ Elektronisches Publizieren und Rechtsdatenbanken, Vorbemerkungen der Herausgeber, in: Schweighofer/Menzel/Kreuzbauer 2001, S. 98.

⁵ Elektronische Publizieren und Rechtsdatenbanken, Vorbemerkungen der Herausgeber, in: Schweighofer/Menzel/Kreuzbauer 2001, S. 98.

den persönlichen Einschätzungen der Systembenutzer fragen, aber keine Evaluierung der tatsächlichen Retrievalleistung vornehmen.⁶

Auch für die in Deutschland größte und wichtigste juristische Online-Datenbank *juris* werden Verbesserungsmöglichkeiten vor allem darin gesehen, weitere intellektuelle Indexierungsmaßnahmen vorzunehmen, z.B. durch die verbesserte Vergabe von Schlagworten und Normen bzw. durch den Aufbau einer juristischen Klassifikation.⁷

Dies ist nicht zuletzt deshalb erstaunlich, weil das System *juris* im Vergleich zu anderen Fachinformationssystemen sowie juristischen Informationssystemen in anderen Ländern⁸ bereits eine starke intellektuelle Indexierungskomponente aufweist. So werden nicht nur die bibliographischen Angaben erfasst, wie z.B. für die Rechtsprechungsdatenbank das Datum, das Gericht, die Gerichtsbarkeit, das Aktenzeichen, die Fundstelle der Entscheidung sowie die Entscheidungen, die die Ausgangsentscheidung zitieren. Daneben erstellen die Dokumentare Kurztexte (die sog. Orientierungssätze), ordnen den Dokumenten die Notation der *juris*-Sachgebietsgliederung zu und vergeben für alle Dokumente Normen und Schlagworte.⁹ Nach der Konzeption von *juris* war die Stichwortsuche, also die Suche mit den aus den Texten mittels des Programms *PASSAT* extrahierten und auf die Grundform zurückgeführten Wörter der Dokumente selbst¹⁰, lediglich dazu gedacht, die Schwächen der intellektuellen Indexierung auszugleichen.¹¹

3.2 Unzuverlässigkeit der Stichwortsuche

Wenngleich die Stichwortsuche bei *juris* letztlich nur als Kompensation der Schwächen der intellektuellen Aufbereitung gedacht war, steht sie heute eindeutig im Vordergrund, die Nutzung der übrigen Register - vielleicht mit

⁶ Vgl. etwa Moritz, BRAK-Mitt 3/1998, S. 115 ff.; ders., Nachrichten für Dokumentation 49 (1998), S. 490 ff.; Riehm et al. 1989, S. 490 ff.

⁷ Etwa Moritz, jur-pc 1995, S. 3000 ff.; Möller 1993, S. 147; Wolf, jur-pc 1992, S. 1744 (1748).

⁸ Hier ist vor allem das weltweit erste als Volltext-Datenbank aufgebaute juristische System *LEXIS* mit Rechtstexten aus 40 Ländern zu nennen, das vor allem in den USA und Großbritannien im Einsatz ist, Möllers, JuS 2000, S. 1203 (1205); Riehm et al. 1989, S. 149.

⁹ Vgl. dazu genauer Mielke 2000, S. 25 ff. und S. 53 ff.

¹⁰ Vgl. dazu genauer Mielke 2000, S. 23 f. und S. 52 f.

¹¹ Vgl. dazu genauer Mielke 2000, S. 64.

Ausnahme des Normregisters - wird allgemein als gering eingeschätzt.¹² Dies ist auch deswegen nicht überraschend, da auch bei der Suche in herkömmlichen Medien überwiegend Schlagwortregister herangezogen werden.¹³

Bei der Stichwortsuche bei *juris* handelt es sich um eine Mischform zwischen Stichwort- und Schlagwortsuche, da die intellektuell vergebenen Schlagworte neben den automatisch extrahierten Stichwörtern in die Recherche miteinbezogen werden.

Gegen die Suche mit Stichwörtern bestehen erhebliche Vorbehalte. So wird grundsätzlich bezweifelt, dass diese Art der Suche für das deutsche Recht geeignet sei, da hier im Gegensatz zum amerikanischen Recht, in dem in weiten Bereichen *case law* vorherrscht, nicht die Suche nach älteren Gerichtsentscheidungen im Vordergrund stehe. Ohne hier in die Einzelheiten zu gehen, sei an dieser Stelle nur darauf hingewiesen, dass auch im deutschen Recht die große praktische Bedeutung von sog. „Präjudizien“ kaum zu bestreiten sein dürfte, wenn auch die herrschende Meinung eine normative Bindungswirkung von Präjudizien ablehnt.¹⁴

Neben diesem rechtsmethodischen Einwand wird die Stichwortsuche vor allem als zu unzuverlässig angesehen. Insbesondere bestünde die Gefahr, dass die Benutzer in großen Datenbanken mit vielen hunderttausend Dokumenten eine nicht mehr handhabbare Menge an Ausgabetexten auf ihre Suchanfrage erhalten.¹⁵ Sucht man z.B. Literatur zu der Rechtsfrage *Besetzungsrüge bzgl. Der Teilnahme eines blinden Richters bei der Augenscheinseinnahme*, erhalte man mit den Suchbegriffen *Richter* und *blind* nicht nur sämtliche prozessualen, gerichtsverfassungsrechtlichen, personalvertretungsrechtlichen, dienst-

¹² Vgl. dazu Kuhlmann, Demokratie und Recht 1988, S. 50 (59 f.); Moritz, BRAK-Mitt 3/1998, S. 115 (117); Wolf, jur-pc 1992, S. 1608 (1612); vgl. auch ausführlich Mielke 2000, S. 65.

¹³ Zu diesem Ergebnis kam eine Untersuchung von Infratest Sozialforschung, Jungjohann et al. 1974, S. 24 f. Vgl. auch Hoffmann 1993, S. 72.

¹⁴ Vgl. dazu ausführlich Mielke 2000, S. 68 f.

¹⁵ Vgl. etwa Moritz, jur-pc 1995, S. 3000 (3002): „Freitextsuche führt meist zu großer Vollständigkeit aber auch zu großem Ballast“; ders., BRAK-Mitt 3/1998, S. 115 (117). Ebenso Bernecker, Verwaltungsrundschau 1994, S. 332 (335); Buhlmann, NJW-CoR 2/1996, S. 98 (99); Schallbruch, Jura 1995, S. 670; Sellnick, ThürVBl 1997, S. 151 (152); Wagner-Döbler, Informatik und Recht 1988, S. 458 (460); Wolf, jur-pc 1992, S. 1608 (1615). Vgl. auch Tenopir/Ro 1990, S. 74 ff. Eine Ausnahme stellt die Untersuchung von Blair/Maron dar, bei deren Studie zu einer englischsprachigen Rechtsdatenbank die precision weitaus höher war als der recall, Blair/Maron, Communications of the ACM 28 (1985), S. 289 (293 ff.).

rechtlichen oder versorgungsrechtlichen Dokumente über blinde Richter, sondern auch solche, in denen sich ein blinder Richter scheiden ließ oder ein sehender Richter von einem Angeklagten entsprechend genannt wurde.¹⁶ Basierend auf dieser Überlegung wird vielfach nahegelegt, dass man mit der Stichwortsuche bevorzugt in Kurztexten und nicht in den Volltexten recherchieren sollte.¹⁷ So bietet die *juris* GmbH in ihrer Benutzeroberfläche *juris formular* als erste Suchmöglichkeit die Suche im Kurztext an, wodurch sich die Suche auf die von Dokumentaren erstellten Kurztexte beschränkt und die Volltexte für die Recherche ausgespart bleiben. Durch diese Gestaltung legt die *juris* GmbH nahe, dass es sich um die erfolgversprechendste Suchmöglichkeit mit Begriffen handelt. Durch eine Einschränkung auf rein formaler Ebene soll damit der Gefahr zu großer Treffermengen begegnet werden.

Den beiden häufig genannten Thesen

- a) die Stichwortsuche führt zu einer zu großen Antwortmenge
- b) im juristischen Bereich ist eine starke intellektuelle Indexierungskomponente erforderlich

soll anhand der Ergebnisse einer Evaluierungsstudie weiter nachgegangen werden. Dabei wurde die Effektivität von Recherchen mittels der Stichwortsuche in *juris* untersucht und mit der Effektivität von Recherchen in einer Datenbank, die statistische Retrievalverfahren nutzt, verglichen.¹⁸ Diese Studie ist dabei dem Umstand zu verdanken, dass ein Teilbereich des Datenbestands, nämlich etwa 400.000 Gerichtsentscheidungen, in beiden Datenbanken in gleichem Umfang gespeichert, aber in unterschiedlicher Weise inhaltlich aufbereitet und damit auf unterschiedliche Art zu recherchieren ist.

4 Evaluierung juristischer Informationssysteme

4.1 Juris vs. QSearch

Bei *juris* handelt es sich um ein „klassisches“ Boolesches Retrievalsystem, bei dem - wie bereits ausgeführt - zur Suche alle sinntragenden Begriffe der gespeicherten Dokumente (Gerichtsentscheidungen, Aufsätze, Gesetze und

¹⁶ Vgl. Wolf, jur-pc 1992, 1608 (1615) sowie ausführlich zu diesem Beispiel Mielke 2000, S. 66 f.

¹⁷ Vgl. etwa Buhlmann, NJW-CoR 2/1996, 98 (99): „Wegen der hohen Trefferzahl eignet sich Volltextsuche meist nicht als alleiniges Suchkriterium“ oder Bettschneider, Versicherungswirtschaft 1998, 550 (551): „Die Suche im Kurztext führt zu exakteren Ergebnissen, daher sollte sie stets bevorzugt vor der Recherche im Gesamttext durchgeführt werden.“ Vgl. dazu auch Mielke 2000, S. 146.

¹⁸ Zum Testaufbau genauer Mielke 2000, S. 105 ff., sowie Mielke 2001, S. 104 ff.

Rechtsverordnungen, Gesetzesmaterialien etc.) in ihrer Grundform sowie die zusätzlich von Dokumentaren intellektuell bzw. manuell vergebenen Suchkriterien zur Verfügung stehen. Bei dem System *QSearch*¹⁹ handelt es sich um ein natürlich-sprachliches System, bei dem der Benutzer mit natürlichsprachlich formulierten Sätzen anfragen kann. Für die Inhaltserschließung kommen statistische Verfahren zum Einsatz, die mit Wortfrequenzen und Termgewichungen arbeiten.²⁰ Eine manuelle oder intellektuelle Aufbereitung der gespeicherten Dokumente ist nicht gegeben. Die Ausgabe erfolgt nach Relevanz geordnet (*ranking*).

4.2 Testanordnung

An der nachfolgend beschriebenen Evaluierungsstudie nahmen insgesamt 22 Doktoranden zum Zivilprozessrecht teil. Die Testpersonen erhielten zwölf Sachverhalte von Entscheidungen zu zivilprozessualen Rechtsproblemen aus den Jahren 1987 bis 1992. Alle Urteile und Beschlüsse sind veröffentlicht und in beiden Datenbanken enthalten. Die in den Entscheidungen veröffentlichten Sachverhaltsangaben bildeten die Aufgabentexte. Dabei ist der Testaufbau nicht so zu verstehen, dass nach einer *einzelnen* vorgegebenen Entscheidung gesucht werden soll. Vielmehr dienten die gewählten Aufgabentexte dazu, ein allgemeines rechtliches Problem - ohne Vorgabe bestimmter Suchbegriffe - aufzuzeigen, zu denen die Testpersonen relevante Entscheidungen suchen sollten. Den Testpersonen war dabei nicht bewusst, dass sich in der Datenbank eine zu den jeweiligen Sachverhalten ergangene Entscheidung befindet.²¹

Als Sachverhalt 12 dienten beispielsweise folgende Angaben:²²

Die Parteien streiten über die Wirksamkeit der Zustellung eines Vollstreckungsbescheides. Die Klägerin macht aus abgetretenem Recht einen Anspruch aufgrund eines Schuldaner-

¹⁹ Diese Datenbank wurde von dem deutschen Recherchedienst der COMPLEX GmbH genutzt, deren Geschäftsanteile seit 1. September 2000 von der Foris AG übernommen wurden. Der Recherchedienst wird dabei weiterhin über Telefon und Fax betrieben, vgl. <http://www.Foris.de>, Stand: 28. Mai 2002. Zum Recherchedienst der COMPLEX GmbH vgl. Mielke 2000, S. 157 ff.

²⁰ Eine auf der US-amerikanischen Patentschrift für dieses System basierende Beschreibung der wichtigsten statistischen Verfahren findet sich in Mielke 2000, S. 154 ff.

²¹ Das Testdesign ist ausführlich beschrieben in Mielke 2000, S. 107 ff.

²² Die Angaben stammen aus dem in BGHZ 111, 1 ff. (=NJW 1990, 1666 f.) veröffentlichten Urteil des VIII. Zivilsenats vom 14. März 1990, Aktenzeichen: VIII ZR 204/89. Alle weiteren Sachverhalte sind in Mielke 2000, S. 188-195 abgedruckt. Eine weitere Übersicht zu den ausgewählten Sachverhalten findet sich in Mielke 2000, S. 197-203.

kenntnisses geltend, durch das der Beklagte nach ihrer Darstellung Kaufpreisforderungen für Warenlieferungen der Zedentin bestätigt hat. Der Beklagte bestreitet sowohl die Warenlieferungen als auch die Echtheit seines Namenszuges unter dem Schuldanerkenntnis. Die Klägerin erwirkte gegen den Beklagten zunächst einen Mahnbescheid und anschließend einen Vollstreckungsbescheid über 8.217,33 DM nebst Zinsen und Kosten. Der Vollstreckungsbescheid wurde am 29. November 1984 in der Wohnung des Beklagten an seine damalige Lebensgefährtin, Frau T., übergeben.

Der Beklagte hatte sich im Jahr 1982 von seiner ersten Ehefrau getrennt und lebt seitdem mit Frau T. und deren minderjährigen Kindern zusammen. Im Dezember 1983 bezog er mit Frau T. und ihren Kindern die Wohnung, in der die umstrittene Zustellung stattfand. Dort führte Frau T. – die der Beklagte zwischenzeitlich geheiratet hat – den gemeinsamen Haushalt, für dessen finanzielle Bedürfnisse der berufstätige Beklagte sorgte.

Der Beklagte hat gegen den Vollstreckungsbescheid, von dessen Existenz er erst aufgrund einer Lohnpfändung erfahren haben will, am 21. Juli 1986 Einspruch eingelegt und in dem anschließenden Verfahren beantragt, den Vollstreckungsbescheid aufzuheben und die Klage abzuweisen. Gegen die Wirksamkeit der Zustellung hat er eingewandt, zum Zustellungszeitpunkt habe zwischen Frau T. und ihm noch keine gefestigte Lebensgemeinschaft bestanden, denn er habe »jederzeit wieder gehen können.« Die Klägerin hat beantragt, den Einspruch als unzulässig zu verwerfen.

Die Testpersonen sollten in einem *paper-and-pencil*-Test die Suchbegriffe angeben, die für eine Datenbankrecherche zu den in den Sachverhalten enthaltenen Rechtsproblemen geeignet erscheinen. Zum voranstehenden Sachverhalt, bei dem es um die Frage geht, ob eine Ersatzzustellung an den nichtehelichen Lebensgefährten des Zustellungsempfängers wirksam ist, konnte man etwa die Suchbegriffe *Zustellung*, *Ersatzzustellung*, *nichteheliche Lebensgemeinschaft* etc. verwenden. Im weiteren Testverlauf dienten die Anfrageformulierungen der Testpersonen als Suchanfragen an die *juris*-Rechtsprechungsdatenbank.²³

Der Test orientiert sich an einer vermittelten Recheresituation, bei der Juristen an professionelle Rechercheure ihre Suchanfragen richten, die dann die eigentliche Datenbankrecherche durchführen. Diese vermittelte Recherche dürfte immer noch in großen Teilbereichen ständige Praxis sein.²⁴ Selbst die *juris*-GmbH bietet für Personen oder Institutionen, die über keinen eigenen *juris*-Zugang verfügen, einen solchen Rechereservice an.²⁵ Auch bei dem Recherchedienst der COMPLEX GmbH, der nun von der Foris AG übernommen worden ist, handelt es sich um eine vermittelte Recherche. Unabhängig davon, ob man diese Art der Recherche für angemessen hält oder nicht, ergibt sich für die Testanordnung somit eine gewisse Praxisnähe. Da offensichtlich

²³ Alle Anfrageformulierungen der Testpersonen finden sich in Mielke 2000, S. 205-216.

²⁴ Siehe dazu auch Mielke 2000, S. 108 f.

²⁵ Vgl. www.juris.de unter Bestellen - Rechercheauftrag, Stand 28. Mai 2002.

fehlerhafte Suchstrategien korrigiert wurden, konnten darüber hinaus technische Probleme oder Schwierigkeiten der Testpersonen mit der Abfragesprache aufgefangen werden.²⁶ Die Testergebnisse sind daher auch als unabhängig von Weiterentwicklungen der *juris*-Schnittstelle seit dem Testzeitpunkt in den Jahren 1995 und 1996 anzusehen.

Als Testmaterial standen 264 Aufgabenlösungen zur Verfügung. Da die Testpersonen zu einigen Aufgaben mehrere Alternativlösungen angeboten haben, ergaben sich insgesamt 312 Suchanfragen, die in einem ersten Schritt als Recherchen in der *juris*-Rechtsprechungsdatenbank dienten. Mit den Anfragen zur Hälfte der zwölf Sachverhalte, insgesamt 152 Suchanfragen, wurde in der *QSearch*-Datenbank recherchiert.

4.3 Testauswertung

Die mit den Suchanfragen der Testpersonen in der *juris*-Rechtsprechungsdatenbank recherchierten Dokumente wurden hinsichtlich der Größe der Antwortmengen und der Relevanz der ausgegebenen Dokumente für den jeweiligen Ausgangssachverhalt analysiert. Dabei wurden insgesamt 1.183 Gerichtsentscheidungen auf der Grundlage des Volltextes intellektuell auf ihre Relevanz für das zugrundeliegende Rechtsproblem untersucht.²⁷

Suchgenauigkeit vs. Wiedergewinnungsrate

Um der These näher nachzugehen, dass die Stichwortsuche insbesondere zu unzuverlässig ist und die Gefahr birgt, zu viele irrelevante Dokumente zu erzielen, wurden zunächst die in der *juris*-Rechtsprechungsdatenbank erzielten Antwortmengen in drei Klassen eingeteilt: Nullantworten, Ausgabemengen in einer akzeptablen Größenordnung sowie Antwortmengen mit zu vielen Dokumenten, da *juris* kein *ranking* der Antwortmenge nach inhaltlichen Kriterien kennt. Fraglich ist, ab welcher Grenze es sich noch um ein – was die Menge der nachgewiesenen Dokumente betrifft – akzeptables Ergebnis handelt. Nach der Literatur und auch nach Auffassung professioneller Rechercheure ist bereits eine Trefferzahl von über 20 Dokumenten zu groß und erfordert weitere Einschränkungen.²⁸ Für die hier vorliegende Studie wurde als

²⁶ Zu den vorgenommenen Korrekturen vgl. Mielke 2000, S. 116-121.

²⁷ Vgl. ausführlich auch zur Methode Mielke 2000, S. 130 ff.

²⁸ Möller 1993, S. 92; Viefhues, NJW-CoR 7/97, S. 343. Vgl. auch Turtle 1994, S. 215. Nach Jordan/Konradi-Martin, AnwBl 1994, S. 117 (119) ist ein Suchergebnis bei einer Treffermenge von fünf bis zehn Entscheidungen optimal. Riehm et. al. 1989, S. 235, gehen ebenfalls davon aus, dass man im Normalfall versucht sein wird, die Antwortmenge auf eine übersichtliche Zahl von 10 bis 20 Dokumenten einzuschränken. Relativ

Grenze für eine handhabbare Menge eine Trefferzahl von 30 Dokumenten angenommen.

Dabei zeigt sich folgendes Ergebnis:

Nullantworten	15,15 %
Treffermengen zwischen 1 und 30	61,36 %
Treffermengen über 30	22,35 %
Entweder Nullantworten oder über 30 Treffer ²⁹	1,14 %

Es zeigt sich damit bereits auf dieser Ebene, dass sich die These, die Stichwortsuche führe zu sehr hohen Antwortmengen, nicht bestätigt. Vielmehr scheinen die Benutzer mit der Stichwortsuche – zumindest was die Anzahl der ausgegebenen Dokumente betrifft – gut zurecht zu kommen, wenn bereits in einem Test, der keine Kontrolle über die ausgegebenen Antwortmengen durch interaktives Recherchieren erlaubt und auf einer Datenbasis von ca. 400.000 Entscheidungen zu allen Rechtsgebieten erfolgte, überwiegend handhabbare Mengen gewonnen werden.

Da Benutzer im Fall akzeptabler Antwortmengen wesentlich seltener ihre Suchstrategie grundlegend ändern werden, als dann, wenn sie extrem hohe Ausgabezahlen oder Nullantworten erhalten, ist davon auszugehen, dass ein Großteil der Probleme der Testpersonen, die bereits mit der ersten Anfrage zu brauchbaren Treffermengen kommen, auch bei einer interaktiven Recherche mit der Möglichkeit einer Strategieänderung auftauchen.

Um eine Aussage zur Qualität der recherchierten Dokumente treffen zu können, wurden insgesamt 1.183 Judikate auf ihre Relevanz hin untersucht. Die Relevanzbewertung beschränkt sich dabei auf die Anfragen mit einer Antwortmenge zwischen 1 und 30 Dokumenten, da davon auszugehen ist, dass sich die Testpersonen eine solche Ausgabe ansehen würden, während sie bei den übrigen Trefferkategorien eine Änderung der Suchstrategie vornehmen würden.

Es zeigt sich dabei, dass die *precision* tendenziell höher ist als der *recall*. Von den ausgewerteten 174 Anfragen, die eine Antwortmenge zwischen 1 und 30 Dokumenten erzielen, wird in insgesamt 67 Fällen (entspricht knapp 39 %)

weit geht Jahnel, jur-pc 1995, S. 3183 (3188), wenn er als verwertbares Endergebnis maximal 50 Fundstellen ansieht. Unabhängig davon mag zu Forschungszwecken das Durchsuchen einer größeren Antwortmenge gerechtfertigt sein.

²⁹ Bezogen auf die verschiedenen Alternativanfragen zu einer Aufgabe.

eine *precision* zwischen 0,5 und 1 erreicht, darunter 43mal (entspricht knapp 25 %) die bestmögliche *precision* 1. Dagegen lässt sich nur für 45 Recherchen (knapp 26 %) ein *recall* von 0,5 und darüber berechnen. Einen *recall* von 1 erreichen lediglich 8 % der Anfragen. Während die *precision* einen Mittelwert von 0,4 erreicht, liegt der *recall* bei durchschnittlich 0,28. Dies bedeutet, dass durchschnittlich 40 % der ermittelten Dokumente relevant sind, gleichzeitig aber nur 28 % aller recherchierbaren relevanten³⁰ Dokumente gefunden werden können. Insgesamt zeigt sich, dass sich die These der mangelnden Genauigkeit als Hauptproblem der Stichwortsuche nicht bestätigen lässt.

Dass der *recall* weiter sinkt, wenn die Suche auf die Kurztexte beschränkt wird, was von einem Großteil der Literatur zur Vermeidung zu hoher Trefferzahlen vorgeschlagen wird³¹, zeigen die Ergebnisse einer Zusatzuntersuchung. Als Testmaterial dienten die hinsichtlich ihres *recall* besten Anfragen jeder Aufgabe, mit denen nun auf den Kurztext beschränkt recherchiert wurde. Als Ergebnis dieses Tests ist festzuhalten, dass sich zwar die Zahl der nicht relevanten Dokumente proportional stärker verringert als die Zahl der relevanten Entscheidungen, dies jedoch mit einem Verlust von ca. 60 % der relevanten Judikate einhergeht. Die Wahrscheinlichkeit, mit der Stichwortsuche relevante Dokumente zu finden, sinkt also erheblich, wenn die Suche auf den Kurztext beschränkt wird.³²

Ergebnisse des Vergleichstests juris – QSearch

Eine Auswertung der ermittelten Dokumentmengen erübrigt sich für die Auswertung des Vergleichstests auf der *QSearch*-Datenbank, da durch das *ranking* das Problem zu großer Antwortmengen entfällt. Ebenso sind aufgrund des *partial match*-Paradigma nur in extremen Ausnahmefällen Nullantworten zu erwarten. Um den gleichen Maßstab wie bei der Auswertung auf der *juris*-Datenbank zu erzielen, bilden die jeweils zuerst angezeigten 30 Dokumente die Grundlage für die Relevanzbewertung.³³ Bewertet man diese Dokumente nach ihrer Relevanz, zeigt sich, dass sich die Zahl der relevanten Dokumente bei der Recherche in der Datenbank mit statistischen Inhaltser-

³⁰ Als Bezugsgröße für die Zahl aller relevanten Entscheidungen diene die Gesamtzahl der von allen Anfragen zu einer Testaufgabe erzielten relevanten Dokumente, vgl. dazu genauer Mielke 2000, S. 95 f. und S. 132. Ein Vergleich mit den in den verschiedenen Kommentaren oder Lehrbüchern zitierten einschlägigen Entscheidungen zeigt, dass der *recall* weiter absänke, wenn man diese mit in die Gesamtzahl einbezüge, da überwiegend nicht alle zitierten einschlägigen Entscheidungen gefunden werden konnten, vgl. dazu genauer Mielke 2000, S. 134 f.

³¹ Vgl. oben Fußnote 17.

³² Zu den genauen Zahlen Mielke 2000, S. 143-145.

³³ Vgl. ausführlich zum Vorgehen Mielke 2000, S. 165 ff.

schließungskomponenten deutlich erhöht. Während die Anfragen bei *juris* nur 175 relevante Dokumente erzielen konnten, waren bei *QSearch* 425 relevante Dokumente zu ermitteln.³⁴

Auch die Zahl der **unterschiedlichen** relevanten Dokumente ist bei *QSearch* höher als bei *juris*. Bei *juris* wurden insgesamt 39 verschiedene relevante Entscheidungen recherchiert, bei *QSearch* 56. Bezieht man die durch die Recherche auf der *QSearch*-Datenbank gewonnenen zusätzlichen relevanten Dokumente in die Gesamtzahl der zu ermittelnden relevanten Dokumente mit ein, verringert sich die Wiedergewinnungsrate bei *juris* auf 20 %. Bei *QSearch* liegt die Rate dann bei 30 %, sie ist damit um die Hälfte höher.

Auch das *ranking* ist als überwiegend zufriedenstellend einzustufen: 80% der Anfragen führen innerhalb der ersten 30 angezeigten Entscheidungen zumindest zu einem relevanten Dokument, wobei sich 56,71 % der relevanten Dokumente bereits innerhalb der ersten zehn angezeigten Dokument befinden.³⁵ *QSearch* gibt damit innerhalb der jeweils ersten zehn angezeigten Dokumente mehr relevante Dokumente aus als *juris* insgesamt.

Insgesamt lassen sich die Ergebnisse wie folgt zusammenstellen:

	<i>Juris</i>	<i>QSearch</i>
Zahl der Anfragen, die die Ausgangsentscheidung finden	50 %	80 %
Gesamtzahl der nachgewiesenen relevanten Dokumente	175	425
Zahl der unterschiedlichen relevanten Dokumente	39	56
<i>Recall</i>	0,2	0,3

5 Fazit

Die Ergebnisse der empirischen Studie zeigen, dass entgegen der Einschätzung der Fachliteratur das Problem der zu großen Antwortmengen bei der Stichwortsuche nicht vorrangig ist. Dies ergibt auch die qualitative Auswertung des *juris*-Tests, wonach die *precision* insgesamt tendenziell höher ist als der *recall*. Damit müssen die Überlegungen zur Verbesserung juristischer Informationssysteme verstärkt dahin gehen, wie der *recall* zu erhöhen ist, was bisher in der Diskussion zu *juris* eine eher geringe Rolle spielt. Auch der Einschätzung vieler Autoren, dass die Suche im Kurztext erfolgversprechender ist, kann vor diesem Hintergrund nicht gefolgt werden.

³⁴ Die Zahlen beziehen sich dabei auf die Anfragen, mit denen in beiden Datenbanken recherchiert wurde.

³⁵ Zu den genauen Zahlen vgl. Mielke 2000, S. 167 und S. 174 ff.

Der Vergleich mit einem Informationssystem, das statistische Verfahren zur Inhaltserschließung nutzt, zeigt, dass die statistischen Verfahren zu einer Erhöhung der Wiedergewinnungsrate führen. Ein natürlichsprachliches System wie *QSearch* ist für den Nutzer damit nicht nur wesentlich komfortabler, sondern stellt auch das leistungsfähigere System dar. Da bei der Stichwortsuche in *juris* auch die intellektuell vergebenen Schlagwörter in die Recherche miteinbezogen werden, ist sogar davon auszugehen, dass der statistische Ansatz diese intellektuelle Indexierungskomponente erfolgreich ersetzt. Eine intellektuelle Inhaltserschließung bringt nur im Bereich der Erfassung formaler Angaben und der Zuordnung von Normen einen Zusatznutzen. Der Test liefert somit auch ein Argument gegen eine gängige Ansicht der deutschen und österreichischen Rechtsinformatik, wonach für Rechtsinformationssysteme intellektuelle Indexierungskomponenten unumgänglich seien. Diese Einschätzung hängt möglicherweise mit der im Vergleich zu anderen Fachgebieten sehr umfangreichen inhaltlichen Erschließung von Rechtstexten in den herkömmlichen Medien zusammen. Es lassen sich aber bislang keine empirisch abgesicherten Hinweise dazu finden, dass die Übertragung dieses Konzepts auf Informationssysteme erfolgversprechend ist.

Auch lässt der Interaktionsmodus von *QSearch* einen mobile access wesentlich realistischer erscheinen als der von *juris*, bei dem nur über ein relativ aufwändig gestaltetes Formular zu recherchieren ist. Damit bietet *QSearch* eine niedrigere Einstiegsschwelle in die Interaktion, die schon heute über einen sprachbasierten, wenn auch bisher nicht automatisierten, Recherchedienst zugänglich ist.

6 Literaturverzeichnis

- Bernecker, H. M.: Das juristische Informationssystem JURIS, Verwaltungsrundschau 1994, S. 332-337.
- Bettschneider, O.: Die Informationsgewinnung mit *juris*. Teil 1: Die *juris* WWW-Oberfläche - Speziallösung für Versicherer, Versicherungswirtschaft 1998, S. 550-553.
- Blair, D. C./Maron, M. E.: An Evaluation of Retrieval Effectiveness for a Full-Text Document-Retrieval System, Communications of the ACM 28 (1985), S. 289-299.
- Buhlmann, D.: Nicht auf Sand gebaut. *Juris* CD-ROM Baurecht im Test, NJW-CoR 2/1996, S. 98-107.
- Hoffmann, H.: Informationstechnik am Richterarbeitsplatz. Gerichts-EDV. Datenbanken. Berechnungsprogramme. Standardsoftware. Entwicklung juristischer Programme, München 1993.

- Jahnel, D.: Elektronische/konventionelle Suchmethoden - ein Gegensatz?, jur-pc 1995, S. 3183-3191.
- Jordan, H./Konradi-Martin, Ch.: Erstattung von juris-Datenbankrecherchekosten. Neue Wege der Rechtsprechung, AnwBl 1994, S. 117-119.
- Jungjohann, K., et al.: Informationsverhalten und Informationsbedarf von Juristen. Eine Erhebung von Infratest Sozialforschung, München, im Auftrag des Bundesministeriums der Justiz und der Gesellschaft für Mathematik und Datenverarbeitung. Teil 1: Analyse-Band, Berlin 1974.
- Kuhlmann, J.: Das juristische Informationssystem JURIS nach der Privatisierung. Auswirkungen auf anwaltliche Chancengleichheit, juristisches Arbeiten und juristisches Denken, Demokratie und Recht 1988, 50 (59 f.);
- Mielke, B.: Bewertung juristischer Informationssysteme. Evaluierung von *juris* im Vergleich zu einem statistischen Information Retrieval-System anhand zivilprozessualer Fragestellungen, Köln, Berlin, Bonn, München 2000.
- Mielke, B.: Wie effektiv sind Recherchen in juristischen Informationssystemen? Ein Vergleich von *juris* mit *QSearch*, in: Schweighofer/Menzel/Kreuzbauer (Hg.), Auf dem Weg zur ePerson, Aktuelle Fragestellungen der Rechtsinformatik 2001, Wien 2001, S. 101-112.
- Möller, T.: *juris* für Juristen. Optimierte *juris*-Nutzung mit den Mitteln der Rechtsinformatik unter Berücksichtigung überkommener juristischer Methodenlehre zugleich ein Beitrag zur Abbildung des deduktiven Hauptschemas der analytischen Begründungslehre als Computermode. Diss. Jur. Saarbrücken 1993.
- Möllers, Th.: Die juristische Recherche in Bibliotheken und mit neuen Informationstechnologien, JuS 2000, S. 1203-1210.
- Moritz, K.: Optimierung der intellektuellen Ordnungskriterien bei *juris* am Beispiel des Arbeitsrechts: Sachgebietsgliederung, Normen, Schlagwörter, jur-pc 1995, 3000 – 3009.
- Moritz, K.: Nutzen eines *juris*-Anschlusses am juristischen Arbeitsplatz, BRAK-Mitt 3/1998, S. 115-119.
- Moritz, K.: Nutzen der juristischen Datenbank „*juris*“ am Richterarbeitsplatz, Nachrichten für Dokumentation 49 (1998), S. 490-492.
- Riehm, U., et al.: Endnutzer und Volltextdatenbanken. Empirische Untersuchungen zur Nutzung von Volltextdatenbanken in den Fachwelten Medizin, Recht und Wirtschaft. Kernforschungszentrum Karlsruhe, Juni 1989.
- Schallbruch, M.: *Juris*-Datenbanken vorgestellt, Teil 1: Die Rechtsprechungsdatenbank, Jura 1995, S. 670.
- Sellnick, H.-J.: Quo vadis juristisches Informationssystem? Ein Szenario zur Nutzung des Internet für Juristen, ThürVBl 1997, S. 151-154.
- Schweighofer, E./Menzel, Th./Kreuzbauer, G. (Hg.): Auf dem Weg zur ePerson, Aktuelle Fragestellungen der Rechtsinformatik 2001, Wien 2001.
- Turtle, H.: Natural Language vs. Boolean Query Evaluation: A Comparison of Retrieval Performance, in: Croft, W. B./van Rijsbergen, C. J. (Hg.): SIGIR '94. Proceedings of the Seventeenth Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval, Dublin City University 1994, S. 213-220.
- Tenopir, C./Ro, J. S.: Full Text Databases, New York et al. 1990.

Viefhues, W.: juris goes WINDOWS, NJW-CoR 7/97, S. 343-345.

Wagner-Döbler, R.: Karlsruher Juristische Bibliographie und juris: Ein Vergleich, Informatik und Recht 1988, S. 458-461.

Wolf, G.: juris – Ein denkbar einfacher Zugang zu allen Informationen, die Sie brauchen? jur-pc 1992, Teil 1: S. 1524-1536; Teil 2: S. 1568-1577; Teil 3: S. 1608-1619; Teil 4: S. 1676-1691; Teil 5: S. 1744-1752; Teil 6: S. 1801-1810.



Ein „Virtuelles Bibliotheksregal“ für die Informationswissenschaft als Anwendungsfall semantischer Heterogenität

Peter Hanke, Thomas Mandl, Christa Womser-Hacker

Informationswissenschaft - Universität Hildesheim
Marienburger Platz 22 - 31141 Hildesheim
hanke_peter@web.de
mandl@uni-hildesheim.de
womser@uni-hildesheim.de

Zusammenfassung

Das virtuelle Bibliotheksregal *MyShelf* soll den Browsing-Zugang zu heterogen repräsentierten Objekten erleichtern. Anwendungsfall sind die informationswissenschaftlichen Literaturbestände an der Universitätsbibliothek Hildesheim. *MyShelf* zielt darauf ab, die relevanten Bibliotheksbestände, Bestände anderer Bibliotheken, von Dozenten erstellte Lehrmaterialien und Quellen aus dem Internet zu integrieren. Der Zugang umfasst mehrere hierarchische Ordnungssystematiken, wobei vor allem bestehende reale Bibliothekssystematiken integriert werden. Die Systematik soll vom Benutzer ausgewählt werden können, woraufhin sich die Bücher nach der gewählten Systematik neu anordnen. Dazu müssen semantische Transferrelationen zwischen den Objekten oder Systematiken definiert werden. Der Beitrag berichtet über die konkreten Schritte zur Erstellung dieses Systems.

Abstract

The virtual library shelf *MyShelf* enhances the browsing access to heterogeneously represented objects. It is applied to the information science books in the library of the University of Hildesheim. *MyShelf* integrates the relevant library stock, other libraries and teaching material in the internet. The access is possible through various hierarchical cataloguing systems and especially library catalogues. The ordering system can be chosen by the user and the system reorganizes its content accordingly. For that purpose, transfer relations between the objects and categories need to be established. This article includes steps toward the implementation of *MyShelf*.



1 Einleitung

Während die semantische Heterogenität im Information Retrieval für Suchsysteme inzwischen weitgehend als Problem erkannt wurde, wird dieser Aspekt beim Browsing-Zugang noch vernachlässigt. Dabei stellt gerade das Vorhandensein heterogener Systematiken, welche sich oft in unterschiedlichen hierarchischen Zugangssystemen äußern, eine sehr augenfällige Folge der Heterogenität dar. Dieser Artikel überträgt und adaptiert die existierenden Lösungsansätze für Suchsysteme auf Informationssysteme, die Browsing-Zugriff erlauben.

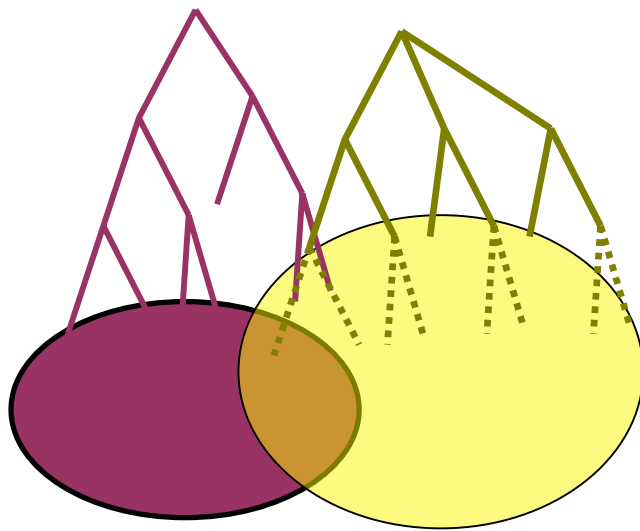


Abb. 1: Heterogene hierarchische Systematiken für sich überlappende Kollektionen

Dazu wird zunächst die Problematik der semantischen Heterogenität und ihrer Behandlung rekapituliert. Anschließend werden die konkreten Anforderungen und Lösungsansätze für das virtuelle Bibliotheksregal *MyShelf* eingeführt, das Browsing in heterogenen Bibliothekskollektionen ermöglicht. Kapitel vier und fünf zeigen im Detail, welche Schritte für die Erstellung einer Datenbasis für die Informationswissenschaft erforderlich waren. Das Fazit diskutiert Möglichkeiten, das System für den Alltagsbetrieb zu realisieren.

2 Semantische Heterogenität

Unerwünschte Heterogenität und die sich daraus ergebende Forderung nach Interoperabilität sind Aspekte der weltweiten Vernetzung von Informationssystemen. Die Heterogenität entsteht dabei v.a. auf drei Ebenen:

- Technische Heterogenität: Informationssysteme benutzen unterschiedliche Hardware, Zeichensätze, Datenformate, Kommunikationsprotokolle u.ä.

- **Metadaten-Heterogenität:** Informationssysteme basieren auf verschiedenen Standards für die Erfassung von Metadaten (Dublin-Core, MARC, usw.) oder unterschiedlichen Datenmodellen.
- **Semantische Heterogenität:** Die ersten beiden Ebenen sind gelöst, die inhaltlichen Repräsentation können technisch verglichen werden, aber die beschreibenden Elemente entstammen unterschiedlichen Ontologien oder Sinnzusammenhängen.

Alle Ebenen erschweren die Kommunikation und damit das Wissensmanagement für die Benutzer.

2.1 Ursachen der Heterogenität

Der Hauptgrund für die semantische Heterogenität liegt in der durchaus wünschenswerten virtuellen Integration von online bereitgestellten Wissensbeständen verschiedenster Anbieter. Die Trennung der Informationsangebote wird aus politischen und inhaltlichen Interessen bestehen bleiben, da viele Informationsanbieter sich durch ihr individuelles Angebot profilieren. Jeder Anbieter setzt für seine Daten seine optimierte Inhaltserschließung ein, ohne auf Kompatibilität mit anderen, möglicherweise konkurrierenden Informationsdienstleistern zu achten. Anbieter wie Fachinformationszentren entwickeln Thesauri für ihre Fachgebiete, die das Vokabular definieren und für die intellektuelle Indexierung verwendet werden. Anbieter mit kleinen Datenmengen greifen zunächst häufig auf eigens gebildete hierarchische Strukturen zu, ohne sich bewusst zu machen, dass hier ein neuer Thesaurus geschaffen wird.

Die semantische Heterogenität besitzt aber noch eine weitere Dimension. Missverständnisse treten nicht nur fachkulturübergreifend auf, sondern sind besonders zwischen verschiedenen nationalen oder ethnischen Kulturen folgenreich. Heterogenität vermindert gegenseitiges Verständnis in der privaten wie auch in der geschäftlichen Kommunikation.

Jede Einteilung der Welt in Begriffe - sei es in Form eines Thesaurus, einer Klassifikation oder einer sonstigen Ontologie - repräsentiert eine meist in sich sinnvolle Perspektive. Diese entspricht aber keineswegs immer der Sichtweise des Benutzers, besonders wenn dieser aus einer anderen Kultur oder auch Fachkultur stammt. Benutzer sind nicht an der globalen Einteilung des verfügbaren Wissensbestandes interessiert, sondern wollen mit den ihnen geläufigen Begriffen suchen.

Verschiedene Begriffssammlungen und Systematiken sind also in aller Regel nicht kompatibel. Es kommt auch oft vor, dass der gleiche Begriff eine andere Bedeutung hat, da er anders in den größeren Kontext eingebettet ist. Analysen solcher Probleme liegen für viele Fachgebiete und Anwendungsfälle vor (cf. z.B. BUCKLAND ET AL. 1999 für Bereiche der Ingenieurwissenschaft, CHEN ET AL. 1998 für die Biologie, MANDL & WOMSER-HACKER 2002 für Lehr- und Lernmaterialien, MANDL 2001 für einen Überblick, einen Überblick über die Rolle von Ontologien im Wissensmanagement bietet STAAB 2002.).

Im Information Retrieval Prozess führt Heterogenität dazu, dass jede Quelle mit den für sie spezifischen Begriffen befragt werden müsste. Aus Sicht des Benutzers stehen der von ihm intendierten Bedeutung des Anfrageterms potentiell mehrere unterschiedliche Bedeutungen gegenüber. Häufig weckt die semantische Heterogenität den Wunsch nach einem weltübergreifenden Ordnungssystem für Wissen. Zahlreiche Versuche in diese Richtung sind bereits unternommen worden und gescheitert. Solche Lösungen sind weder durchsetzbar, noch könnten sie die erforderliche Dynamik für die Integration neuen Wissens bieten. Ideen dieser Art stehen hinter dem sogenannten *Ontology Merging* (cf. NOY & MUSEN 1999).

Bei gemischten Korpora, die sowohl intellektuell als auch automatisch indexierte Dokumente enthalten, verschärft sich die semantische Heterogenität. Diese Situation liegt etwa im Internet vor, wo Suchmaschinen automatisch indexieren, aber auch intellektuelle Inhaltsanalysen in Form von Meta-Daten des Autors oder einer Zuordnung etwa in einem Subject Guide vorliegen können. Die Möglichkeiten der Virtualisierung stellen das Information Retrieval also vor die Herausforderung, Suchen in semantisch stark heterogenen Umgebungen zu ermöglichen.

2.2 Behandlung der Heterogenität

Die Behandlung von Heterogenität erfolgt über Transferrelationen zwischen verschiedenen Ontologien. Kuhlen spricht in diesem Zusammenhang von einer „postkoordinierenden Ordnung durch transformierende Anpassung“ (KUHLEN 1999:138). Zwei Familien von Methoden können dafür eingesetzt werden (cf. HELLWEG ET AL. 2001):

- Intellektuell definierte Übergänge zwischen Ordnungssystematiken
- Assoziative, vage Beziehungen, die durch maschinelles Lernen abgeleitet werden (cf. MANDL 2001)

Erstere führen zu sicheren Beziehungen, die etwa in Form von wenn-dann-Regeln formalisiert werden.

2.3 Maschinelles Lernen

Die Grundlage für vage Transfer-Relationen stellen maschinelle Lernverfahren dar. Maschinelles Lernen beschäftigt sich mit Software, welche aufgrund von Eingaben ihr Verhalten bzw. ihren Output verändern kann. Die Konstruktion eines Modells des Gegenstandsbereichs erlaubt es, Aufgaben besser zu erledigen. Das maschinelle Lernen unterscheidet verschiedene Lernstrategien:

- Lernen durch Deduktion
- Lernen durch Analogie
- Induktives Lernen (anhand von Beispielen)

Alle diese Strategien lassen sich auch für Transferrelationen einsetzen. Vor allem kommt der induktive Ansatz zur Anwendung, bei dem meist aus vielen Beispielen ein quantitatives Modell abgeleitet wird.

Vage Relationen lassen sich als vage Regeln darstellen, wie etwa in Form von Wahrscheinlichkeiten für bestimmte Beziehungen. Grundlage für die vagen Verfahren sind Objekte, bei denen die Beziehungen zu allen Systematiken bekannt sind bzw. schon intellektuell erschlossen wurden. Damit bildet gewissermaßen auch hier die menschliche Informationsarbeit die Basis. Daneben kommen aber auch Heuristiken für die Ableitung von Doppelkorpora zum Einsatz (cf. MANDL 1999). Maschinelle Lernverfahren extrapolieren die Beziehungen dann auf andere Objekte. Besonders wichtig sind in diesem Zusammenhang die Erfahrungen aus der Text-Kategorisierung (cf. JOACHIMS 2002). Da die Auswahl des für den Anwendungsfall besten maschinellen Lernverfahrens ein Problem darstellt, kommen zunehmend Kombinationen von mehreren Algorithmen zum Einsatz (cf. MANDL & WOMSER-HACKER 2001). Während die Erstellung intellektueller Übergänge sehr teuer ist, läuft sie bei vagen Relationen automatisch ab.

2.4 Anwendungen

Die zahlreichen Anwendungen von Systemen zur Heterogenitätsbehandlung ergeben sich aus den Anforderungen und finden sich in allen Bereichen, in denen integrierte Informationssysteme auf Objekte mit mehreren unterschiedlichen Repräsentationsgrundlagen zugreifen.

Wichtige Anwendungen zeichnen sich derzeit im Umfeld der Vision des semantic web ab, das durch intensives Markup von Web-Seiten datenbankähnli-

che Abfragen im Internet ermöglichen will (cf. BERNERS-LEE ET AL. 2001). Ein Beispiel bieten DOAN ET AL. 2002, die eine Abbildung zwischen den im Internet zugänglichen Kursen verschiedener Universitäten erstellen. Dies erfordert das Mapping heterogener Fachbezeichnungen, Niveaus und Inhalte aufeinander.

3 Das virtuelle Bibliotheksregal *MyShelf*

Semantische Heterogenität überwiegt trotz zahlreicher Standardisierungsbemühungen nach wie vor auch im Bibliotheksbereich.

3.1 Ausgangssituation

Die Präsenz-Bibliothek hat auch im Zeitalter Digitaler Bibliotheken ihre wichtige Rolle noch nicht verloren. Der Gang zum Regal repräsentiert gewissermaßen den Urahn aller Browsing-Systeme und stellt häufig eine adäquate Strategie bei der Literatursuche dar. Die Vorteile liegen neben dem direkten Zugriff in dem Ordnungssystem. Die dem Benutzer vertraute lokale Aufstellungssystematik liefert einen erheblichen Mehrwert. Der direkte Zugriff bietet weitere Vorteile. Manche Benutzer erkennen anhand der Gestaltung des Buches den Verlag und können eine erste Relevanzentscheidung treffen. Auch der schnell ersichtliche Umfang kann bei bestimmten Informationsbedürfnissen eine Rolle spielen.

Der Nachteil hierarchischer Systematiken wie auch des Bibliotheksregals liegt in der Starrheit. Das Ordnungssystem repräsentiert immer nur eine Sichtweise von vielen möglichen, was auch zu einer Vielzahl von Bibliotheksklassifikationen geführt hat. Auf aktuelle Entwicklungen innerhalb der Universitäten oder der Wissenschaften kann die Bibliothek nur langsam reagieren. Dabei würde sie die angestrebte Konsistenz gefährden. Das Bibliotheksregal ist vom Umfang her beschränkt, es fehlen Titel aus den nicht frei zugänglichen Beständen, ebenso gibt es oft keinen Hinweis auf ausgeliehene Titel oder Bücher in Semesterapparaten (die wiederum eine neue Sichtweise auf die Dokumente einführen). Äußerst schwierig ist meist das Einfügen von Dokumenten, die nicht vom Typ „Buch“ sind. Video- und Audio-Dokumente sowie Datenträger werden oft getrennt aufbewahrt und das Einstellen von Unterlagen für Lehrveranstaltungen, die von Dozenten selbst erstellt wurden, ist meist nicht vorgesehen.

Die Literaturversorgung für das kürzlich eingerichtete Fach Informationswissenschaft an der Universität Hildesheim bietet einen Anwendungsfall, bei

dem zahlreiche Aspekte der semantischen Heterogenität auftreten. Die Literaturversorgung und Ausstattung der Bibliothek ist besonders aus studentischer Sicht nicht optimal, da keine Signatur für die Informationswissenschaft existiert und relevante Literatur an sehr unterschiedlichen Orten aufgestellt wird. Die Studierenden greifen aufgrund des notwendigerweise beschränkten Umfangs auf viele andere Quellen zu und werden dort mit anderen Systematiken konfrontiert.

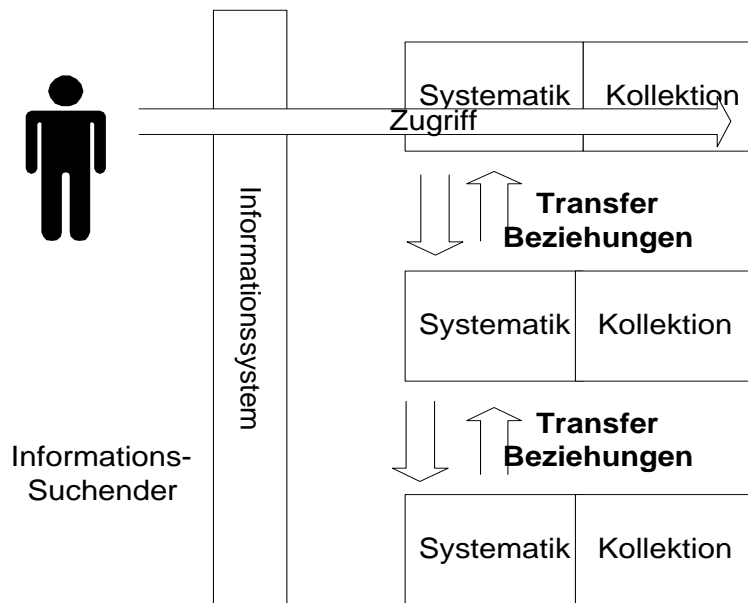


Abbildung 2: Die Lösung der Heterogenität in MyShelf: Transfer-Beziehungen erlauben den Zugriff auf mehrere Kollektionen über eine Ontologie

Die Informationswissenschaft ist stark interdisziplinär ausgerichtet, daher sind viele Perspektiven auf das Gebiet möglich, was sich in den Aufstellungssystematiken informationswissenschaftlicher Literatur an verschiedenen Universitäten in Deutschland widerspiegelt.

Quellen aus dem Internet und elektronische Kursmaterialien gewinnen zwar zunehmend an Bedeutung und Relevanz für die Ausbildung, sie sind aber nicht nach bibliothekarischen Gesichtspunkten erschlossen und nicht am gleichen Standort zu finden wie relevante Bücher. Wissenschaftler und Studierende wünschen sich, von ihrem Arbeitsplatz aus auf die relevanten Materialien zuzugreifen.

3.2 Lösungsansatz *MyShelf*

Das Ziel des virtuellen Bibliotheksregals *MyShelf* besteht in einer virtuellen Integration von mehreren Dokument-Beständen, bei der sich die semantische

Heterogenität nicht nachteilig für den Anwender auswirkt. Zunächst kann der Benutzer also mit *MyShelf* auf den maximal verfügbaren Korpus und damit auf alle Objekte zugreifen. Darüber hinaus bietet das System auch alle Systematiken an, welche die einzelnen Teilkorpora beschreiben. Im Anwendungsfall Informationswissenschaft sollen also möglichst viele relevante heterogene Bibliothekssystematiken eingehen. Damit hat jeder Benutzer die ihm vertraute Systematik zur Verfügung, er kann aber auch auf andere Sichtweisen ausweichen, um zu prüfen, ob er bei einem konkreten Suchproblem mit einer anderen Systematik eventuell erfolgreicher ist.

Der entscheidende Mehrwert besteht aber darin, dass immer die gesamte Kollektion nach der aktuell ausgewählten Systematik geordnet wird. Das virtuelle Bibliotheksregal kennt also mehrere Organisationsmöglichkeiten für seinen Inhalt und passt diese den Präferenzen des Benutzers an.

3.3 Benutzungsoberfläche

Die Präsentation und die Interaktion des Benutzers mit großen Mengen von Kategorien stellt nach wie vor ein Problem dar, das für jeden Anwendungsfall erneut optimiert werden muss. Allein die Existenz zahlreicher Alternativen für die Darstellung und Bedienung von hierarchischen Strukturen zeigt die Vielfalt der Gestaltungsmöglichkeiten und bei genauer Analyse zahlreiche Detailschwierigkeiten (cf. BEKAVAC 1999). Die für den *MyShelf*-Ansatz entscheidende Visualisierung mehrerer hierarchischer Strukturen wurde bisher selten realisiert. MUKHERJEA ET AL. 1995 präsentieren eine Anwendung in der für Hypertext-Strukturen mehrere Ursprungsknoten gewählt werden, die jeweils in eine andere Hierarchie münden. Die parallele Präsentation mehrerer Produktthesauri in einem wirtschaftlichen Informationssystem findet sich bei MANDL & STEMPFHUBER 1998.

Für die Darstellung nicht hierarchisch sondern assoziativ verbundener Konzepte entstehen zunehmend stärker visualisierte Darstellungsformen wie zweidimensionale Karten (cf. EIBL & MANDL 2002). Diese könnten sich besonders für die Darstellung von vagen Transferrelationen eignen, die durch Verfahren des maschinellen Lernens abgeleitet wurden. In diesem Beitrag steht die Visualisierung nicht im Fokus.

4 Bewertung von Bibliotheken

Der Zugriff auf informationswissenschaftliche Bücher in der Universitätsbibliothek Hildesheim stellt, wie bereits aufgezeigt, einen konkreten Anwendungsfall für das Problem der semantischen Heterogenität dar. Um Studierenden und Wissenschaftlern einen leichteren Zugang zur gesuchten Literatur zu ermöglichen, sollte exemplarisch eine einheitliche Sichtweise auf den informationswissenschaftlichen Bücherbestand der UB Hildesheim erarbeitet werden. Zu diesem Zweck musste die entsprechende Literatur erfasst werden, um sie anschließend einer neuen Ordnungssystematik zuzuführen.

Zur Erschließung des relevanten Buchbestandes in der UB Hildesheim wurde der Weg des Abgleichs mit dem Literaturangebot anderer Bibliotheken beschritten. Dabei boten sich Standorte informationswissenschaftlicher Studiengänge an. Für die Durchführung wurden die Bibliotheken der Universitäten Konstanz, Regensburg und Saarbrücken ausgewählt. Um Aussagen darüber treffen zu können, wie geeignet die einzelnen informationswissenschaftlichen Bestände für einen Abgleich mit den Titeln der Hildesheimer UB sind, erfolgte eine stichpunktartige Bewertung der genannten Bibliotheken.

Im Rahmen einer Literaturrecherche mit 200 verschiedenen Titeln der Informationswissenschaft sollte aufgezeigt werden, welche Überschneidungen zum Hildesheimer Bestand existieren. Jeweils 50 der zu recherchierenden Titel stammten dabei aus den Universitätsbibliotheken Konstanz, Regensburg, Saarbrücken bzw. Hildesheim. Sie wurden aus 20 unterschiedlichen Systemstellen der dortigen informationswissenschaftlichen Klassifikationen entnommen (abgesehen von Hildesheim, wo Schwerpunktbereiche der dortigen Lehre gewählt wurden).

Die Auswertung der Recherchen zeigte, dass Konstanz die größten Überlappungen mit den Hildesheimer Beständen aufweist (cf. HANKE 2002:36 ff.). Von den insgesamt 80 in Hildesheim gefundenen Titeln wurden 56 auch in Konstanz geführt. In Regensburg belief sich die Schnittmenge auf 48 Titel, von diesen waren jedoch nur fünf in der Regensburger Verbundklassifikation unter einer informationswissenschaftlichen Signatur erfasst oder mit einer entsprechenden Nebenstelle versehen; dies trifft ebenfalls auf nur 34 der insgesamt 110 gefundenen Regensburger Titel zu. In Konstanz hingegen besaßen von den 56 Titeln 41 eine Signatur der dortigen Ordnungssystematik „kid“ (Kybernetik, Informatik, Datenverarbeitung und Informationswissenschaft), so dass ein Abgleich mit den Daten einer Bestandsdatei des „kid“ die besten Ergebnisse versprach. In Saarbrücken fanden sich im Vergleich dazu nur 26

der 80 Hildesheimer Titel wieder, von denen bis auf fünf auch alle im „kid“ geführt werden. Die Gesamtbestände der in den einzelnen Bibliotheken unter einer informationswissenschaftlichen Signatur eingeordneten Bestände unterstützen die Ergebnisse der Studie: in Konstanz sind es mehr als 23.000, in Regensburg weniger als 1.000 und in Saarbrücken etwa 2.500 Titel (Stand: April 2002). Der im Folgenden beschriebene Abgleich ergab für die UB Hildesheim eine Menge von ca. 5.000 informationswissenschaftlichen Titeln.

5 Extrahieren des informationswissenschaftlichen Buchbestandes

Um den informationswissenschaftlichen Bücherbestand der Universitätsbibliothek möglichst umfassend zu bestimmen, sollte ein automatischer Abgleich mit einer geeigneten Bestandsliste von Titeln durchgeführt werden. Aus den beschriebenen Gründen wurde dafür die Bestandsdatei des „kid“ ausgewählt. Das entsprechende Gegenstück, eine Datei mit relevanten Hildesheimer Titeln, musste erstellt werden. Eine Erfassung des Gesamtbestandes der Hildesheimer UB kam aus technischen Gründen nicht in Betracht.

Somit wurde alternativ die Ermittlung einer relevanten Teilmenge anvisiert, welche anhand einer Heuristik erarbeitet wurde. Dazu wurde eine Liste informationswissenschaftlicher Begriffe erstellt, die sich am so genannten „alphabetischen Register“ der Universitäts- und Landesbibliothek Düsseldorf orientierte¹. Eingang fanden auch Bezeichnungen von Systemstellen der Bibliotheksklassifikationen in Konstanz, Regensburg und Saarbrücken, und Klassifikationsbeschreibungen der UB Hildesheim, sowie Begriffe der Basisklassifikation.

Diese Schlüsselbegriffe dienten der Bestimmung von Notationen in der UB Hildesheim, die informationswissenschaftliche Aspekte beinhalten. Dafür wurden sämtliche Klassifikationsbeschreibungen der UB Hildesheim nach den Begriffen durchsucht. Nachdem nun diese Notationen bekannt waren, konnten die zugehörigen Titel per OPAC-Suchanfragen nach den entsprechenden Systemstellen heruntergeladen werden. Die Gesamtmenge der so gefundenen Titel repräsentierte die Bestandsdatei der informationswissenschaftlichen Titel der UB Hildesheim. Denkbar wäre auch gewesen, direkt mit den Schlüsselbegriffen Suchanfragen an den OPAC zu richten, ohne den Umweg über die Notationen zu wählen. Es hatte sich aber im Verlauf der Untersuchungen gezeigt, dass Titelstichwort- und Schlagwortsuchanfragen keine

¹ <http://www.uni-duesseldorf.de/WWW/ulb/iudreg.html>.

zufriedenstellenden Ergebnisse erwarten lassen. Dies liegt sowohl an der teils mangelnden Aussagekraft von Buchtiteln, als auch an der nicht vollständigen bzw. dem Benutzer nicht vertrauten Verschlagwortung (cf. HANKE 2002:24 ff.).

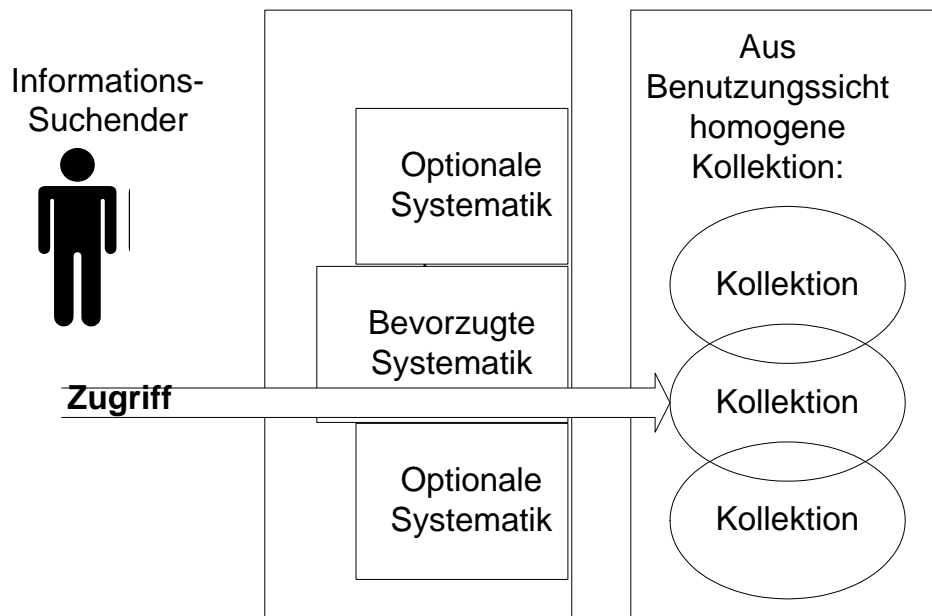


Abbildung 3: Benutzersicht auf MyShelf

Genau wie die vorliegende Bestandsdatei des „kid“ wurden die Hildesheimer Titel geparkt, so dass beide Dateien am Ende eine einheitliche Struktur aufwiesen: pro Titel eine Zeile mit dem Autor, einem Kurztitel, sowie in der Hildesheimer Datei der Signatur und zugeordneten Nebenstellen (Notationen). Die so bearbeiteten Dateien wurden automatisch abgeglichen. Um die Identität von Büchern zu erkennen, mussten Heuristiken entwickelt werden. Ergänzend dazu wurden, um beim Abgleich nicht erfasste Titel zu ermitteln, Literaturanschaffungen von Dozenten informationswissenschaftlicher Lehrveranstaltungen berücksichtigt. Auch Suchanfragen mit Bezeichnungen einer neu erstellten „virtuellen“ Ordnungssystematik für Hildesheimer Bücher der Informationswissenschaft erweiterten den Kreis der Titel. Diese neue Systematik für Hildesheim basiert im Wesentlichen auf den Inhalten der Klassifikationen der Bibliotheken in Konstanz, Regensburg und Saarbrücken, sowie auf der vom Hochschulverband für Informationswissenschaft (HI e.V.) aufgestellten Gliederung für informationswissenschaftliche Lehrveranstaltungen. Die Organisation der neuen Systematik orientiert sich an qualitativen Analysen der anderen Bibliothekssystematiken und versucht, den Zugang vor allem für Studierende zu erleichtern.

Der Ordnungssystematik wurden die beim Abgleich ermittelten Hildesheimer Titel teils halbautomatisch, über die Liste der Schlüsselbegriffe mit zugeordneten Notationen, teils intellektuell zugeordnet.

Die Hildesheimer Studierenden sollen so die Möglichkeit erhalten, unabhängig von der bestehenden lokalen Aufstellungssystematik, über diese neue Ontologie auf informationswissenschaftliche Bücher zugreifen zu können. Durch die Berücksichtigung weiterer informationswissenschaftlicher Systematiken wird die Erstellung intellektuell definierter Transferrelationen stark erleichtert.

Ein weiteres Ziel wäre, über den Einsatz von lernenden Verfahren zu erreichen, dass automatisch erstellte Transferrelationen zwischen den bestehenden Klassifikationen der verschiedenen Bibliotheken erstellt werden. Die im Abgleich zwischen den Bibliotheken gefundene Datenmenge ließe sich für diese Aufgabe anpassen. Sie bilden einen Doppelkorpus, da darin für die gleichen Bücher je eine Repräsentation in zwei heterogenen Systematiken vorliegt. Damit lässt sich ein virtuelles Bibliotheksregal realisieren, das sich dadurch auszeichnet, dass der Bestand mehrerer Bibliotheken zur Verfügung steht, dass nach wie vor alle Systematiken eingesetzt werden und der Benutzer jederzeit zwischen Systematiken wechseln kann.

6 Ausblick

Damit die virtuelle Ordnungssystematik und deren Inhalte zur Anwendung gebracht werden können, müssen Realisierungskonzepte mit der Universitätsbibliothek erarbeitet werden. Dabei wäre vorstellbar, dass die beim Abgleich ermittelten Titel über ihre Signaturen den einzelnen Systemstellen der neuen Klassifikation zugeordnet werden. Beim Anwählen eines Teilgebietes der Klassifikation würde dann eine Suchanfrage nach den entsprechenden Signaturen gestartet. Neue Titel würden automatisiert eingestellt, indem bei bestimmten Suchanfragen (die genauer zu bestimmen wären) die Signaturen der Treffer automatisch einer Systemstelle zugeordnet würden. Auf diese Weise wären keine Einträge in die Datenbank notwendig.

Alternativ dazu könnten neue Bücher bei der Einstellung durch die Bibliothekare eine informationswissenschaftliche Nebenstelle (Notation) zugewiesen bekommen, so dass sie dann über diese bei einer Abfrage auffindig gemacht werden könnten. Dabei müssten die Bibliothekare jedes neue Buch, das sie in die Klassifikation einordnen, mit der passenden Notation versorgen. Zu die-

sem Zweck wäre es notwendig, vorab die gesamte neue Ordnungssystematik mit einer zu bestimmenden Notation auszuzeichnen. Die informationswissenschaftliche Klassifikation würde somit zum Bestandteil der bestehenden lokalen Systematik der UB Hildesheim.

Eine Realisierung außerhalb des Bibliothekssystems in einer eigenen Schicht erlaubt auch die spätere Integration von Internet-Dokumenten. In einem Prototyp von MyShelf werden die neue Systematik und das virtuelle Bibliothekssystem mit Studierenden evaluiert.

7 Literaturverzeichnis

- Bekavac, Bernard (1999): Suche und Orientierung im WWW. Verbesserung bisheriger Verfahren durch Einbindung hypertextspezifischer Informationen. Konstanz: Universitätsverlag [Schriften zur Informationswissenschaft Bd. 37].
- Berners-Lee, Tim; Hendler, James; Lassila, Ora (2001): The Semantic Web. A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. In: Scientific American. 17. Mai 2001.
<http://www.sciam.com/issue.cfm?issueDate=May-01>
- Buckland, Michael; Gey, Fred; et al. (1999): Mapping Entry Vocabulary to Unfamiliar Metadata Vocabularies. In: D-Lib Magazine vol. 5(1).
<http://www.dlib.org/dlib/january99/buckland/01buckland.html>
- Chen, Hsinchun; Martinez, Joanne; Kirchhoff, Amy; Ng, Tobun; Schatz, Bruce (1998): Alleviating Search Uncertainty through Concept Associations: Automatic Indexing, Co-Occurrence Analysis, and Parallel Computing. In: Journal of the American Society for Information Science. JASIS 49(3). S. 206-216.
- Doan, AnHai; Madhavan, Jayant; Domingos, Pedro; Haelvy, Alon (2002): Learning to Map between Ontologies on the Semantic Web. In: Proceedings of the (WWW2002)
<http://www2002.org/CDROM/refereed/232/index.html>
- Eibl, Maximilian; Mandl, Thomas (2002): Including User Strategies in the Evaluation of Graphic Design Interfaces for Browsing Documents. In: Skala, Vaclav (ed.): Journal of WSCG. Special issue. The 10th Intl Conf in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG 2002). Pilsen. vol 1. S. 163-169.
http://wscg.zcu.cz/wscg2002/Papers_2002/B89.pdf
- Hanke, Peter (2002): Neue Chancen und Möglichkeiten für Ordnungssystematiken durch Virtualisierung: Anwendung am Beispiel der Erfassung und Klassifizierung des informationswissenschaftlichen Bücherbestands der Universität Hildesheim. Masterarbeit.
- Hellweg, Heiko; Krause, Jürgen; Mandl, Thomas; Marx, Jutta; Müller, Matthias N.O.; Mutschke, Peter; Strötgen, Robert (2001): Treatment of Semantic Heterogeneity in Information Retrieval. IZ-Arbeitsbericht Nr. 23, IZ Sozialwissenschaften, Bonn.
http://www.gesis.org/Publikationen/Berichte/IZ_Arbeitsberichte/index.htm#ab23
- Joachims, T. (2002): Learning Text Classifiers with Support Vector Machines. Kluwer.

- Kuhlen, Rainer (1999): Die Konsequenzen von Informationsassistenten: Was bedeutet informationelle Autonomie oder wie kann Vertrauen in elektronische Dienste in offenen Informationsmärkten gesichert werden? Frankfurt a.M.
- Mandl, Thomas (1999): Effiziente Implementierung von statistischen Assoziationen im Text-Retrieval. In: Ockenfeld, Marlies; Mantwill, Gerhard (eds.): Information und Region. 51. Jahrestagung der Deutschen Gesellschaft für Informationswissenschaft und Informationspraxis e.V. (DGI). Hamburg, 21.-23.9.1999. S. 159-172.
- Mandl, Thomas (2001): Tolerantes Information Retrieval: Neuronale Netze zur Erhöhung der Adaptivität und Flexibilität bei der Informationssuche. Konstanz: Universitätsverlag [Schriften zur Informationswissenschaft Bd. 39].
- Mandl, Thomas; Womser-Hacker, Christa (2001): Fusion Approaches for Mappings Between Heterogeneous Ontologies. In: Constantopoulos, Panos; Sølvberg, Ingeborg (Hrsg.): Research and Advanced Technology for Digital Libraries: 5th European Conference (ECDL 2001) Darmstadt 4.-8.9.2001. Springer [LNCS 2163]. S. 83-94.
- Mandl, Thomas; Stempfhuber, Maximilian (1998): Softwareergonomische Gestaltung von Wirtschaftsinformationssystemen am Beispiel von ELVIRA. In: Ockenfeld, Marlies; Mantwill, Gerhard (eds.): Information und Märkte. Kongreß der Deutschen Gesellschaft für Dokumentation (DGD). Bonn, 22.-24.9.98. Frankfurt am Main. S. 145-157.
- Mandl, Thomas; Womser-Hacker, Christa (2002): Anforderungen an Information Retrieval Systeme in Bildungsmärkten. In: Schmidt, Ralph (Hrsg.): Content in Context: Perspektiven der Informationsdienstleistung. 24. Online-Tagung der DGI. Frankfurt am Main. 4-6. Juni 2002. S. 67-78.
- Mukherjea, Sougata; Foley, James; Hudson, Scott (1995): Visualizing Complex Hypermedia Networks through Multiple Hierarchical Views. In: Proc ACM Conference on Human Factors and Computing Systems (SIGCHI 1995) Denver. Colorado. http://www.acm.org/sigchi/chi95/Electronic/documnts/papers/sm_bdy.htm
- Noy N.; Musen, M. (1999): SMART: Automated Support for Ontology Merging and Alignment. In: Twelfth Workshop on Knowledge Acquisition, Modeling and Management. Banff, Alberta, Canada. 16.-21. Oktober, 1999. <http://sern.ucalgary.ca/KSI/KAW/KAW99/papers/Fridman1/NoyMusen.pdf>
- Staab, Steffen (2002): Wissensmanagement mit Ontologien und MeataDaten. In: Informatik Spektrum 25(3). S. 194-209.



In: Hammwöhner, Rainer; Wolff, Christian; Womser-Hacker, Christa (Hg.): Information und Mobilität, Optimierung und Vermeidung von Mobilität durch Information. Proceedings des 8. Internationalen Symposiums für Informationswissenschaft (ISI 2002), Regensburg, 8. – 11. Oktober 2002. Konstanz: UVK Verlagsgesellschaft mbH, 2002. S. 303 – 321

ENFORUM - Stand und Aussichten

*Bernard Bekavac, Joachim Griesbaum, Rainer Kuhlen,
Andreas Lenich, Thomas Schütz, Wolfgang Semar*

Universität Konstanz
Fachbereich Informatik und Informationswissenschaft
Fach D87
78457 Konstanz

{Bernard.Bekavac, Joachim.Griesbaum, Rainer.Kuhlen, Andreas.Lenich,
Thomas.Schuetz, Wolfgang.Semar}@uni-konstanz.de

Zusammenfassung

Das von der Konstanzer Informationswissenschaft initiierte Projekt ENFORUM¹ (Enzyklopädisches Forum - Förderung im Rahmen von GlobalInfo) hat seit Anfang März 2002 das in Zusammenarbeit mit dem Paderborner Groupware Competence Center (GCC) entwickelte elektronische Wörterbuch (www.enforum.net) in den routinemäßigen Betrieb genommen.

ENFORUM ist ein internetbasiertes virtuelles Wörterbuch mit enzyklopädischen Eigenschaften auf dem weiteren Gebiet der Informationswissenschaft, welches sich intensiv auf die Hypertextmethodologie und die Potenziale elektronischer Kommunikationsforen stützt. In Abgrenzung zu den sonst im Internet verfügbaren Online-Wörterbüchern bzw. Enzyklopädien werden bei ENFORUM Begriffseinträge nicht von einzelnen Autoren oder einem zentralen Redaktionsteam erstellt, sondern kollaborativ erarbeitet und in den jeweiligen Anwendungskontext gestellt. Dabei wird, ausgehend von einem konzeptuellen Vorschlag, über moderierte elektronische Kommunikationsforen ein Begriff von der an ENFORUM beteiligten Scientific Community mit dem Ziel diskutiert, einen inhaltlichen Konsens bezüglich einer endgültigen Begriffsdefinition zu finden.

Die genannte Scientific Community ist keineswegs eine von vornherein fest definierte Benutzergruppe, vielmehr sind alle informations-/dokumentations-

¹ Das Konzept von ENFORUM wurde von Rainer Kuhlen im Rahmen einer Präsentation auf der ISI'00 in Darmstadt vorgestellt.



Dieses Dokument wird unter folgender [creative commons](http://creativecommons.org/licenses/by-nc-nd/2.0/de/) Lizenz veröffentlicht:
<http://creativecommons.org/licenses/by-nc-nd/2.0/de/>

wissenschaftlichen Gruppen an Hochschulen, Instituten und in der freien Wirtschaft (zunächst auf den deutschsprachigen Bereich beschränkt) angesprochen und zur Teilnahme an ENFORUM (www.enforum.net) eingeladen. Aus den bisher beteiligten Gruppen in Deutschland, Österreich und der Schweiz haben sich einige Kompetenzzentren zu speziellen Teilgebieten der ENFORUMs-Terminologie gebildet, welche sich bereit erklären, konzeptuelle Vorschläge und notwendige Begleitaufgaben, allen voran Moderationsfunktionen bei den Diskussionen, zu übernehmen.

1 Einleitung

1.1 Virtuelles Forum

ENFORUM ist ein „virtuelles Enzyklopädisches Wissensforum“, das bei der Begriffsfestlegung auf eine kommunikative Komponente setzt, mit deren Hilfe unterschiedliche virtuelle Redaktionsgruppen die Begriffe und ihre Inhalte im Diskurs festlegen. Mit ENFORUM wurde somit eine innovative Konzeption einer virtuellen Enzyklopädie am Beispiel des Terminologiebestandes des weiteren Informations- und Kommunikationsgebietes verwirklicht.

1.2 Wissensplattform

ENFORUM als virtuelle begriffsorientierte Wissensplattform (Forum) ist ein gemeinsames Unternehmen der (zunächst deutschsprachigen) Scientific Community im weiteren Informationsgebiet unter der Federführung der Informationswissenschaft respektive des Hochschulverbandes für Informationswissenschaft (HI) und unter der koordinierenden Leitung des Lehrstuhls für Informationswissenschaft an der Universität Konstanz (Prof. Kuhlen). Aber auch Arbeitsbereiche der Informatik, Wirtschaftsinformatik, Kommunikations-, Medien- und Bibliothekswissenschaft sind einbezogen.

1.3 Verteilte Organisation

Die Besonderheit dieses Projekts besteht darin, dass es keine statischen Beiträge gibt, die zentral z. B. von einem Redaktionsteam erstellt werden. Vielmehr sind die Einträge in ständiger „Bewegung“. ENFORUM leistet durch seine verteilte Organisation und durch den Einsatz von elektronischen Kommunikationsforen, die zur Erarbeitung der Wörterbucheinträge als Arbeitsmittel genutzt werden, somit auch einen Beitrag zum besseren Verständnis der

Formen kooperativen Wissensmanagements und des elektronischen Publizierens.

Zwar gelten auch in ENFORUM - nicht zuletzt, um in die Ausbildung eine gewisse terminologische Sicherheit zu bringen - Einträge als „fertig“, wenn die Fachwelt über sie einen Konsens hergestellt hat, aber bis dahin und im Prinzip auch danach können alle Beiträge ständig von allen Beteiligten kommentiert, diskutiert, annotiert, bewertet und ggf. modifiziert werden. Einträge sind so natürlich *real da*, aber in der Dynamik und durch die verteilte Organisation *virtuell*.

1.4 Orientierung im Wortschatz des Informationsgebietes

Neben der Erstellung eines wachsenden und aktuellen Terminologiebestandes wird auf diese Art eine umfassende Diskussion im Fachgebiet kontinuierlich geführt. Dies führt zu einer Orientierung im Wortschatz des weiteren Informationsgebietes (Informationswissenschaft, Dokumentation, Archiv, Bibliothek, angrenzende Gebiete wie Informatik, Wirtschaftsinformatik, Kommunikationswissenschaft etc.) – zum Nutzen der die Terminologie gebrauchenden Fachöffentlichkeit, aber auch der allgemeinen Öffentlichkeit, die durch die alle Bereiche durchdringenden Dienste des Internet bzw. durch den allgemeinen Vorgang der Informatisierung mit dieser Terminologie konfrontiert wird.

1.5 Methodische und softwaretechnische Realisierung

Die technische Realisierung und die koordinierende Redaktionsarbeit werden von der Konstanzer Arbeitsgruppe geleistet. Die ENFORUM-Software wurde auf Basis von Lotus Notes unter Betreuung des Groupware Competence Center (GCC - Prof. Ludwig Nastansky, Universität Paderborn) entwickelt.

2 Aktueller Stand von ENFORUM

Die aktuelle ENFORUM-Software (www.enforum.net) ist aufgeteilt in eine Autorenkomponente, eine Benutzerverwaltung und eine Informations- bzw. Präsentationskomponente. Diese ermöglichen zum einen die Eingabe und die Diskussion von Begriffsvorschlägen² und zu anderem unterstützen sie den Anwender durch Präsentations-, Interaktions- und Such-/Navigationsfunktionen bei der Arbeit mit ENFORUM bzw. bei der Erarbeitung von konzeptio-

² Eine ausführliche Beschreibung des organisatorischen Ablaufs innerhalb von ENFORUM findet sich in (Semar et al. 2002).

nellem Wissen. Im Nachfolgenden sollen die einzelnen Komponenten detaillierter beschrieben und jeweils Ideen und Konzepte für zukünftige Erweiterungen gegeben werden.

2.1 Autorenkomponente

Die Autorenkomponente dient der Eingabe von ENFORUM-Einträgen und zugehörigen Diskussionsbeiträgen. Dabei werden zurzeit zwei systematische Einträge unterschieden: Begriffskandidaten und Konzeptvorschläge.

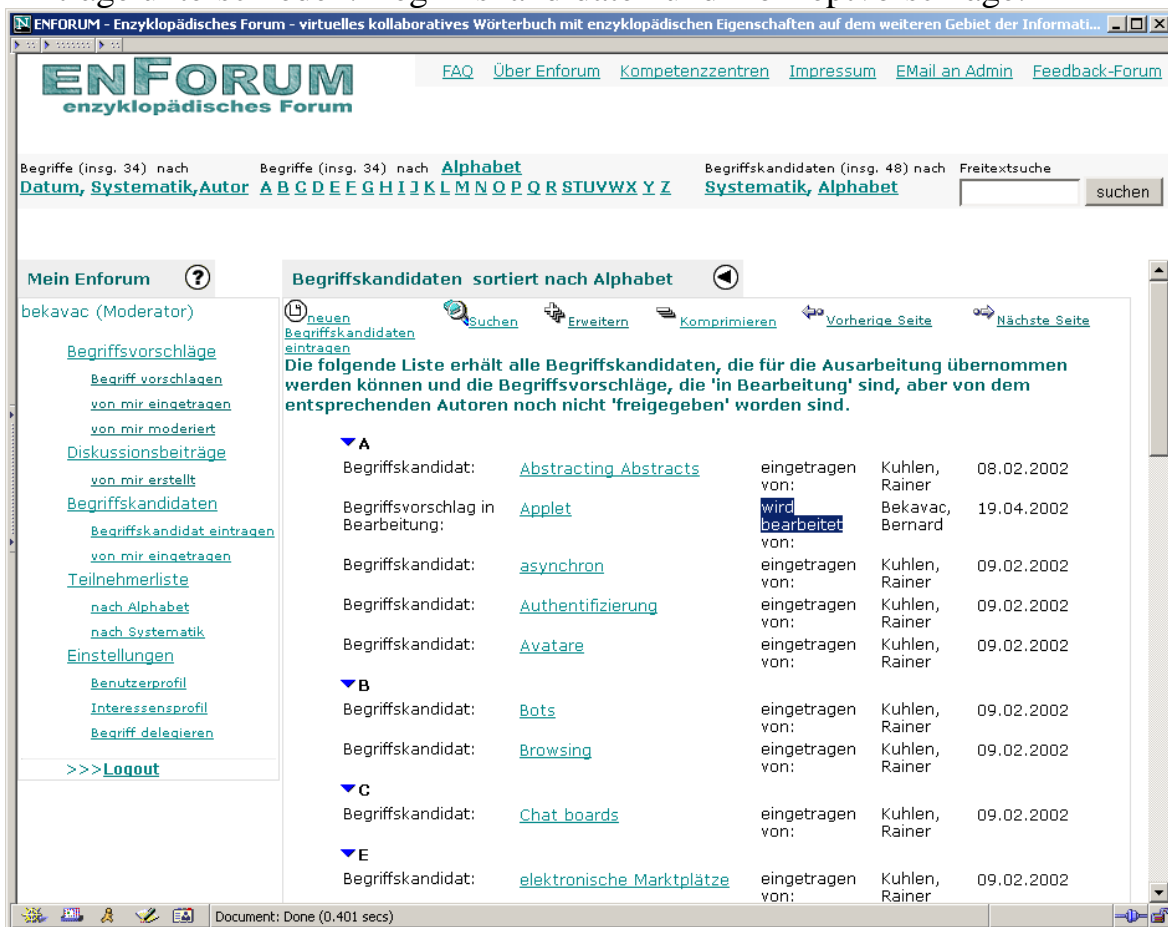


Abb. 1: Liste der aktuellen Begriffskandidaten

Jeder aktive und eingetragene Nutzer von ENFORUM hat das Recht, neue Begriffe, für die noch kein konzeptueller Vorschlag vorliegt, in Form von Begriffskandidaten zur Aufnahme in ENFORUM vorzuschlagen. Ob ein Kandidat tatsächlich dann in ENFORUM aufgenommen und somit ein konzeptueller Vorschlag dazu gemacht werden soll, entscheidet die Fachöffentlichkeit. Entsprechende Abstimmungsverfahren sind allerdings noch Forschungsgegenstand und in der aktuellen Version nicht eingebaut.

Aus der so entstehenden Liste von Begriffskandidaten kann dann ein geeigneter und sich dafür zuständig fühlender Konzeptautor einen Begriff auswählen und für diesen einen Vorschlag erarbeiten. Eine eventuelle gleichzeitige Bearbeitung eines Begriffs durch mehrere Autoren wird dabei durch folgenden Mechanismus vermieden: Sobald ein Begriffskandidat zur Bearbeitung ausgewählt wurde und so lange für diese noch kein konzeptueller Vorschlag vorliegt, wird dieser in der Liste als „in Bearbeitung“ gekennzeichnet (Abb. 1).

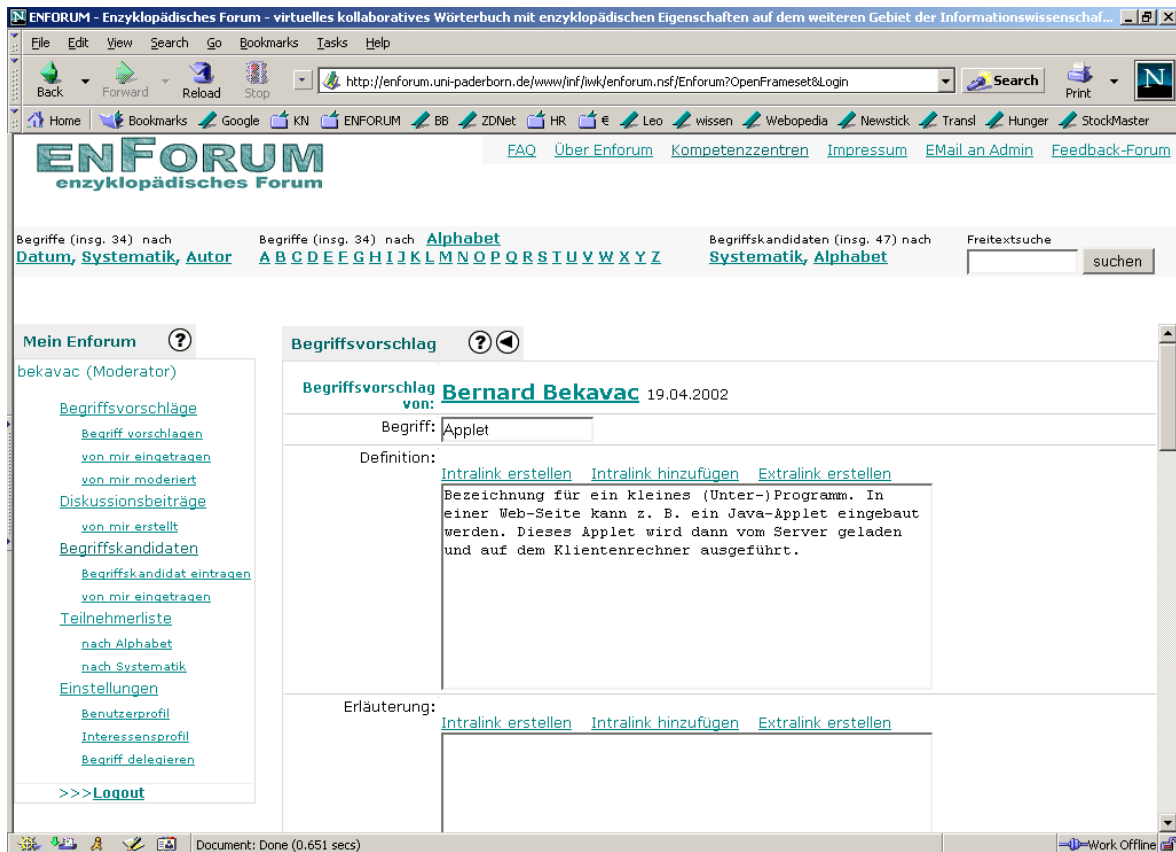


Abb. 2: Erstellen bzw. Bearbeiten eines konzeptuellen Vorschlags für einen Begriffskandidaten³

Wie schon angedeutet, ist die Ausarbeitung eines konzeptuellen Vorschlags, im Gegensatz zu dem Kandidatenvorschlag, nur facheinschlägigen Autoren vorbehalten. Dies entspricht zunächst der Vorgehensweise bei klassischen Wörterbüchern/Enzyklopädien. Wie die Bezeichnung „Vorschlag“ jedoch schon erahnen lässt, wird das ausgearbeitete Konzept dann nicht einfach als Glossareintrag übernommen, sondern über ein zugeordnetes moderiertes Kommunikationsforum in der Fachöffentlichkeit breit diskutiert. Das Vor-

³ Ein konzeptueller Vorschlag ist in ENFORUM ausschließlich nur für den Autoren sichtbar, solange dieser ihn für die Öffentlichkeit nicht freigibt. Somit kann ein Autor einen Begriffsvorschlag über mehrere ENFORUM-Sitzungen hinweg bearbeiten.

schlagen eines neuen Begriffs in ENFORUM erfolgt über vier strukturierte Grundelemente: Begriff (Gegenstand des Eintrags), Definition (eine Real-, Operational- bzw. Nominaldefinition, in der der Inhalt des Begriffs in kurzer und prägnanter Form festgelegt wird), Erläuterung (systematischer und begriffsgeschichtlicher Artikel mit ausführlicher Referenzierung auf die tatsächliche Verwendung sowie die Einbeziehung von verwandten bzw. abgrenzenden Begriffen) und die systematische Einordnung des Begriffs in die im Projekt entwickelte und von der Fachwelt akzeptierte Systematik informationswissenschaftlicher Hauptgruppen (siehe Tab. 1). Optional können relevante Literaturstellen über ein einfaches Erfassungsschema eingetragen werden, die dem Leser die Möglichkeit geben, sich vertiefend in das jeweilige Thema einzuarbeiten. Ebenso ist ein Upload von Artikeln, Grafiken oder sonstigen Materialien möglich, die sich bereits mit dem Begriff beschäftigen. Verknüpfungen zu externen Quellen können direkt in den Wörterbucheinträgen angebracht und so aus dem aktuellen Kontext heraus verstanden und genutzt werden. Abb. 2 zeigt die Erstellung eines Konzeptuellen Vorschlags für einen Begriffskandidaten (s. dazu auch Abb. 1).

Ein weiterer Bereich der Autorenkomponente ermöglicht die Eingabe von Kommentaren zu den konzeptuellen Vorschlägen über ein Kommunikationsforum: Sobald ein Autor seinen Begriffsvorschlag für die Öffentlichkeit freigibt, wird dies den ENFORUM-Nutzern mittels einer Push-Funktion bekannt gegeben. Zu jedem Begriff wird mit seiner Freigabe automatisch ein Diskussionsforum (Konzeptforum) generiert, das von Moderatoren geleitet wird, die auf einzelne Fachgebiete spezialisiert sind. Über die Relevanz und Einschlägigkeit eines Begriffsvorschlags entscheidet so die Fachöffentlichkeit, die durch die Moderatoren dazu aufgefordert wird, an den Diskussionen teilzunehmen. Die Aufgabe der Moderatoren besteht weiterhin in der Initiierung, Leitung und Zusammenfassung der Diskussion sowie der Präsentation der Ergebnisse in ENFORUM. Konzeptforen dienen aber auch der Sammlung von Materialien zur laufenden Veränderung bzw. Erweiterung des Konzeptvorschlags und sind damit ein wichtiger Bestandteil eines disziplinären Wissensmanagements.

Die Erarbeitung der Begriffsdefinitionen über die Kommunikationsforen ist zunächst auf die Verwendung der deutschen Sprache eingeschränkt, um auch den Studierenden den Zugang und die Mitwirkung zu erleichtern. Durch Diskussion und Moderation der Begriffseinträge bewirkt ENFORUM nicht nur die terminologische Kontrolle des Fachgebietes sondern nähert sich auch den Leistungen einer elektronischen Zeitschrift des Informationsgebietes an. Alle Beiträge im ENFORUM werden namentlich gekennzeichnet; die Autoren

werden über Kurzbiographien (und Verweise zu ihren vollständigeren Homepages) ausgezeichnet.

Die Öffentlichkeit hat freien lesenden Zugriff zu den Foren und zu den Begriffseinträgen, nur die registrierte Fachöffentlichkeit hat aktives Schreibrecht.

2.2 Benutzerverwaltung

Die Benutzerverwaltung kontrolliert die Regelung des Lese-/Schreibzugriffs in ENFORUM über fest vorgegebene Rollen. Diese sind hierarchisch angeordnet und folgendermaßen aufgebaut:

- *Gast* (hat nur Lesezugriff) - entspricht im weiteren Sinne der „Öffentlichkeit“
- *IW-Friend* (erhält zusätzlich das Schreibrecht bei den Diskussionsforen)
- *IW-Community* (kann zusätzlich Begriffsvorschläge vornehmen)
- *Moderator* (erhält zusätzlich Editierrechte für bestimmte Bereiche)
- *Administrator* (volle Schreib-/Leserechte, Vergabe von Rollen)

Bei der Anmeldung eines neuen Benutzers wird von diesem ein Rollenwunsch angegeben. Per Default wird dann zunächst die Rolle „Gast“ zugewiesen und später, nach Prüfung des Rollenwunsches durch die Konstanzer Koordinationsgruppe, die gewünschte Rolle vergeben. Ferner werden in der Benutzerverwaltung alle weiteren benutzerspezifischen Angaben (Benutzerprofil) und Einstellungen (Interessensprofil, z. B. für den Push-Dienst) gespeichert.

2.3 Informations- / Präsentationskomponente

Die Informations- / Präsentationskomponente ermöglicht den Benutzern den individuellen Zugriff auf Begriffsvorschläge und die zugehörigen Diskussionsbeiträge auf der Basis diverser Explorations- und Suchmöglichkeiten. Zur weiteren Orientierung im ENFORUM wird, entsprechend dem im System weitgehend verfolgten Hypertextansatz, den Benutzern Gelegenheit gegeben, neben verschiedenen Such- und Navigationsmöglichkeiten, in dem komplexen Begriffsgeflecht auf graphisch anspruchsvollem Niveau zu navigieren, um die Nachbarschaft des interessierenden Begriffs explorieren zu können. Dabei stehen folgende Zugriffsformen in ENFORUM zur Verfügung:

- Systematischer und alphabetischer Zugriff auf Begriffskandidaten mit der jeweiligen Anzeige entsprechender Metaangaben (Initiator, Datum der Eingabe, Bearbeitungsstatus etc.)

- Zugriff auf ein bestimmtes Konzept, sortiert nach Datum, Alphabet, Autorennamen oder nach Kategorien, mit der jeweiligen Anzeige des Begriffsvorschlags und den entsprechenden Metaangaben (Autor, Concept Moderator, Datum des Ersteintrags, Verweis auf Versionen etc.)
- Freitextsuche nach Konzeptvorkommen in den Volltexten, differenzierbar nach Begriffsvorschlag und Diskussionskommentar. Unterstützt wird die Freitextsuche durch diverse Suchhilfen (Boole'sche-, Fuzzy-, Proximity-Suche, Wildcards usw.), die innerhalb der Systemumgebung des Lotus-Domino-Servers zur Verfügung stehen (s. Abb. 3).
- Direkter Zugriff auf die vom (aktuellen) Benutzer eingetragenen Begriffskandidaten, Begriffsvorschläge und Diskussionsbeiträge über eine personalisierte Benutzerleiste (s. „Mein Enforum“ im linken Bereich bei allen Abbildungen). Ferner enthält diese auch den Zugriff auf die persönlichen Options- und Profilangaben (Einstellung des Email-Push-Dienstes etc.)

Zur Minimierung des „Lost in Hyperspace“-Phänomens wurde in ENFORUM auch eine grafische Orientierungs- und Navigationshilfe implementiert (Lenich 2002). Diese ist als zweidimensionale Darstellung von Dokumenten im Umfeld des aktuell betrachteten Begriffskonzepts und ihre organisatorische Einordnung in der zugehörigen Begriffsvernetzung innerhalb von ENFORUM ausgelegt (s. Abb. 4). Durch die kontextabhängig mit der fortlaufenden Navigation aufgebaute Visualisierung werden gleichermaßen eine Orientierung bezüglich des aktuellen Standorts im hypertextuellen Umfeld wie auch Hilfe für die Navigation zu weiteren Inhalten und damit eine Reduzierung der kognitiven Belastung für den Benutzer gegeben.

Es existiert in Bezug auf den aktuellen Begriffseintrag oder Diskussionskommentar eine Ansicht mit Schwerpunkt auf Dokumenten, die über Hyperlinks referenziert werden bzw. auf das aktuelle Dokument verweisen. Eine weitere Ansicht ermöglicht die Einordnung des Begriffs in die Systematik der informationswissenschaftlichen Hauptgruppen von ENFORUM auf graphischer Ebene. Schließlich erfolgt eine Wiedergabe semantisch ähnlicher Dokumente auch, wenn diese über keine hierarchische oder assoziative Verbindung zum gegebenen Dokument verfügen. Die Fokussierung dieser Navigationshilfe auf das aktive Dokument bewirkt eine Komplexitätsreduktion gegenüber den oben beschriebenen alphabetischen oder systematischen Zugriffsformen, da diese, wegen ihrer Unabhängigkeit vom Anwendungskontext, nur wenig zur Orientierung des Benutzers und der Erarbeitung eines mentalen Modells beitragen können und sich eher für den zielgerechten Zugriff auf ein bestimmtes Dokument eignen. Gerade für ENFORUM als eine Anwendung

mit einer in Zukunft erwarteten großen Anzahl informationeller Einheiten kann eine durch explizite Mittel der Komplexitätsreduktion ergänzte Visualisierung eine für den Benutzer mit geringem mentalen Aufwand verbundene Orientierungsform realisieren.



Abb. 3: Freitextsuche bei ENFORUM

Neben der Orientierungsfunktion gibt es in dieser graphischen Navigationsunterstützung auch die Möglichkeit der direkten Navigation durch Anklicken der mittels Icons symbolisierten Dokumente. Die unterschiedlichen Dokumentarten, wie Begriffseintrag, Diskussionskommentar oder ein externes Dokument werden zur visuellen Differenzierung mit unterschiedlichen Icons gekennzeichnet.

Zusätzlich zu den navigationsrelevanten Angaben werden auch Metaangaben (Erstelldatum, Datum des letzten Zugriffs, Autor, Anzahl bisheriger Zugriffe usw.) für das betreffende Dokument und die nachgewiesenen Dokumente der Umgebung zur Verfügung gestellt.

So kann das Umfeld des momentan interessierenden Begriffs oder Forenbeitrags aus verschiedenen Sichten exploriert werden und die Navigation direkt

aus der Orientierungs- und Navigationshilfe zu weiteren Dokumenten verzweigen. Die Orientierungs- und Navigationshilfe bietet damit eine Ergänzung zu der systematischen und alphabetischen Zugriffsform von ENFORUM.

Durch eine periodische Aktualisierung der Metadaten für die grafische Orientierungs- und Navigationshilfe wird ein Kompromiss zwischen einer konsequenten Anpassung der Metadaten infolge neuer oder veränderter Dokumente und einem für die Anpassung des mentalen Modells des Benutzers erforderlichen "sanften" Übergang ermöglicht.

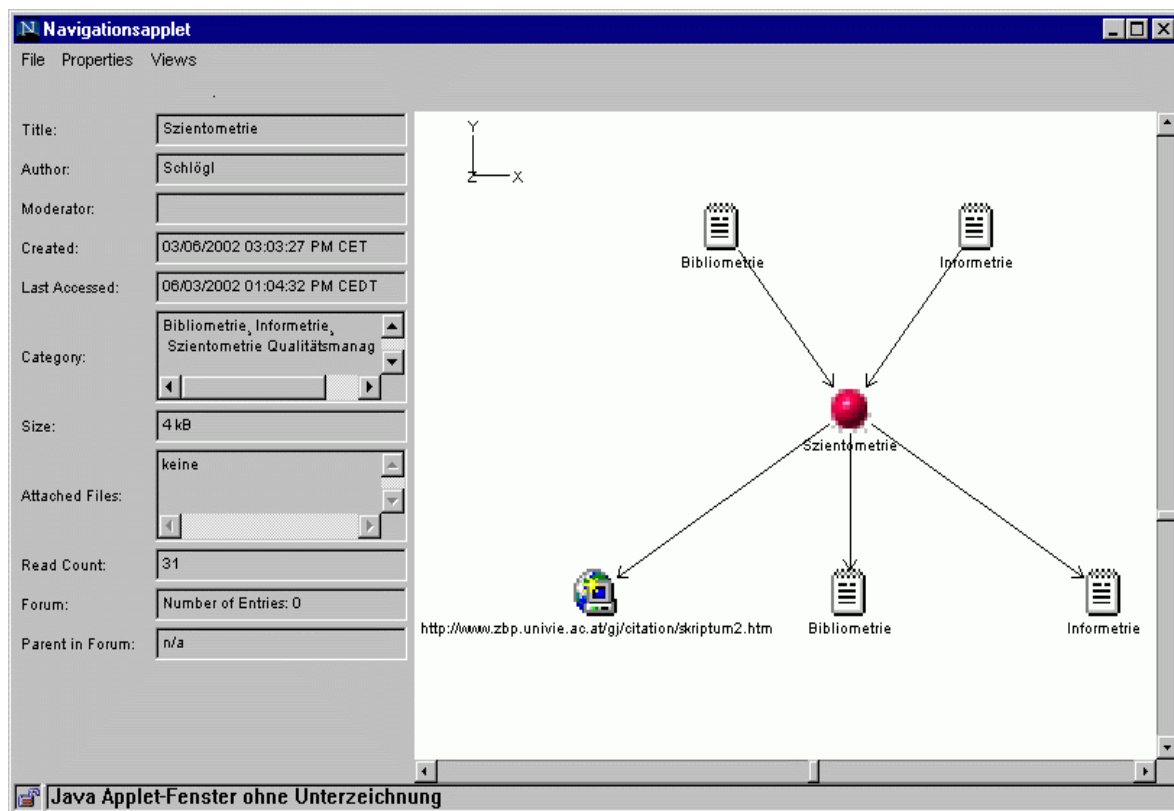


Abb. 4: Grafische Orientierungs- und Navigationshilfe in ENFORUM

Mit der Erkenntnis, dass es nicht die eine Visualisierung nach dem Prinzip "one-size-fits-all" für unterschiedliche Aufgaben und unterschiedliche Nutzer mit entsprechend heterogenen Fähigkeiten und Bedürfnissen gibt, wurde die Orientierungs- und Navigationshilfe als ein System mit potenziell mehreren Visualisierungsformen für dieselben navigationsrelevanten Metadaten konzipiert. Das wird umgesetzt, indem die serverseitige Aufbereitung der Metadaten unabhängig von der Visualisierung beim Client durchgeführt wird und die aufbereiteten Metadaten mittels Extensible Stylesheet Language (XSL) in ein beliebiges Datenaustauschformat transformiert werden können, welches die

Anforderungen der jeweiligen Visualisierungskomponente in Bezug auf ein Eingabeformat erfüllt.

3 Weiterentwicklung und Aussichten von ENFORUM

3.1 Glossareintrag

Für die Zukunft sieht das ENFORUM-Konzept noch eine dritte Art von Begriffseinträgen vor, die Glossareinträge bzw. -definitionen. Ein konzeptueller Vorschlag soll in einen (vorerst) endgültigen Glossareintrag überführt werden, wenn die Diskussion zu dem Vorschlag „zu Ende“ geführt wurde und somit in der Fachöffentlichkeit ein Konsens bezüglich der Definition und Anwendung des Begriffs erreicht ist. Vorerst endgültig bedeutet, dass der Begriff bei Bedarf wieder durch eine erneute Diskussion in einen konzeptuellen Vorschlag „zurückgestuft“ und dann weiterentwickelt werden kann. Noch zu klären sind dabei die Fragen, wann ein Konzeptvorschlag tatsächlich „zu Ende“ diskutiert ist, d.h. wann ein Konsens erreicht ist, und wann eine Glossardefinition nicht mehr ausreichend ist und der Status „endgültig“ aufgehoben werden muss. Nach bisherigen Erfahrungen bedeutet es keinesfalls, dass bei einem Vorschlag ein Konsens erreicht ist, wenn die zugehörige Diskussion zum Erliegen kommt oder wenn erst gar keine entsteht. Dafür kann es sicherlich viele andere Gründe geben, nicht zuletzt kann es an der fehlenden „kritischen Masse“ an facheinschlägigen ENFORUM-Teilnehmern oder an wenig motivierender Moderation liegen.

Naheliegend für die Fragen, wann ein konzeptueller Vorschlag aufgrund seiner Diskussion endgültig geklärt ist und wann eine erneute Diskussion aufgenommen werden kann bzw. darf, wäre die Delegation der Entscheidung an den Moderator. Dieser könnte zu einem Begriffskonzept bestimmen, wann aufgrund des zugehörigen Diskursverlaufs keine neuen Erkenntnisse zu erwarten sind, und die Diskussion dann für beendet erklären. Die zentralen Erkenntnisse aus der Diskussion könnten dann mit dem vorliegenden konzeptuellen Vorschlag abgestimmt werden und so der vorerst endgültige Glossareintrag zustande kommen. Ebenso könnte der Bedarf nach einer neuen Diskussion direkt an den Moderator gerichtet werden und dieser nach Abwägung der Argumente, evtl. unter Rücksprache mit früheren Diskussionspartnern zu dem Begriff, die Neuaufnahme der Diskussion einleiten.

Zur Unterstützung der Moderatorfunktion, auch damit die Entscheidungskraft des Moderators nicht zu subjektiv wird, wird eine zusätzliche Abstimmungskomponente in Erwägung gezogen, mit der dann im demokratischen Sinne

über die Überführung eines Konzeptvorschlags in einen Glossareintrag bzw. die entsprechende Rückstufung, durch die ENFORUM-Teilnehmer abgestimmt werden kann.

3.2 Honorierung (Crediting/Rating) und Abrechnung

ENFORUM soll in weiterer Perspektive ein differenziertes (experimentelles) Honorierungssystem zur Anrechnung der Leistung der Beitragenden erhalten, einschließlich der Lösung der Urheberrechts-/Copyrightfragen, sowie ein ebenfalls differenziertes Abrechnungssystem und Finanzierungsverfahren mit Blick auf eine mögliche kommerzielle Verwertung. Basis der Anrechnung für erbrachte Leistungen ist ein Management-System, das Buch darüber führt, welche Leistung von wem erbracht wurde. Kriterien für die Bewertung/Gewichtung von Beiträgen in den Foren können beispielsweise die Reaktionszeit auf das Einbringen eines Beitrags, der Umfang des Beitrags, die Anzahl der mit dem Kommentar verbundenen Verknüpfungen oder auch die Gewichtung der quantifizierbaren Werte aufgrund des erreichten Expertenstatus des Beitragenden sein. Bei der Einschätzung des Wertes eines Beitrags könnte die folgende Hierarchie bestehen:

- Verknüpfungseintrag
- Reaktion auf einen existierenden Forumsbeitrag
- Verknüpfung mit Bewertung
- Originalkommentar in einem Forum
- Verknüpfung mit Beschreibung der Ziel-Web-Site
- Glossareintrag
- Monitoring-Leistung eines Moderators (CM)
- Anfertigen eines systematischen Konzeptartikels

Im Rahmen einer späteren Zusammenarbeit mit einem kommerziellen Partner kann das Crediting-System als Inverses für ein Abrechnungssystem auf der Basis des Pricing for information (also der realen Nutzung) verwendet werden. Hier sind faire Kompromisse zwischen dem Interesse der Wissenschaft an freier und freizügiger Nutzung von ENFORUM und dem Interesse eines Wirtschaftspartners an kommerzieller (direkter oder indirekter) Verwertung zu entwickeln.

3.3 Suche und Navigation

Im Bereich der gezielten Suche über Stichworteingaben entspricht ENFORUM in seiner aktuellen Version dem State-of-the-Art bei internetbasierten

Wörterbuchsystemen. Die aktuelle Freitextsuche und die zur Verfügung stehenden Suchoperatoren (Möglichkeit von Suchraumeingrenzung usw.) sollen aber noch um ein Passagenretrieval erweitert werden, d.h. an jeder Position, sei es der Glossar-, Forumseinträge oder der systematischen Konzeptartikel, sollen Passagen im Text markiert werden können, zu denen dann ähnliche im System gesucht und nachgewiesen werden.

Besonderen Stellenwert bei der Weiterentwicklung der Suchkomponente sollen jedoch mehr hypertextgerechte Verfahren bekommen. Hierbei soll die Vernetzungs- und Organisationsstruktur von Begriffseinträgen/Diskussionskommentaren untereinander bzw. der Systematik mit in die Suche einbezogen werden. Ziel ist es, auf diese Weise von dem bisherigen reinen Matching-Paradigma wegzukommen und den Nachweis mehrerer bezüglich einer Fragestellung miteinander in Relation stehender ENFORUM-Einträge zu ermöglichen (vgl. Bekavac 1999). So kann eine kontextuelle Einbettung der Einzeltreffer erreicht werden, die durch entsprechende grafische Visualisierungstechniken effizient aufbereitet werden kann.

Weitere Ideen zur Verbesserung der Suchkomponente in ENFORUM zielen auch auf externe Ressourcen ab. Dabei wird nicht nur an die im Internet frei verfügbaren Quellen (andere Wörterbuchsysteme usw.) gedacht, sondern auch an eine Schnittstelle zu Host-basierten Wissensdatenbanken. Bei einer gegebenen Fragestellung könnten damit, neben dem Nachweis von Ergebnissen innerhalb von ENFORUM, automatisch auch relevante Informationen aus externen Quellen extrahiert und in der Ergebnispräsentation von ENFORUM eingebunden werden. Die automatische Einbindung externer Informationen (d.h. nicht nur Verweise, sondern auch Inhalte) ist sicherlich auch generell zu überlegen (vgl. Abschnitt 3.4). Jedes Begriffskonzept würde dann, automatisch und mit einer periodischen Aktualisierung, mit relevanten externen Informationen angereichert, die dann bei der Betrachtung des Konzepts durch den Benutzer mit angezeigt werden würden. Durch die zunehmende Verwendung semantischer Strukturierungsmöglichkeiten im Internet (XML, Dublin Core, RDF, INDECS/ DOI) wird auch hier eine Erleichterung im Verfahren und eine höhere Qualität erwartet.

Im Bereich der Navigationsunterstützung wird neben der Weiterentwicklung des unter 2.3 beschriebenen grafischen Navigationstools auch an eine andersartige Visualisierung des thematischen Umfeldes eines Begriffs gedacht. Vorbild hierzu kann die entsprechende Darstellung innerhalb des Internet-Lexikon „wissen.de“ (www.wissen.de) dienen. Das dort als Matrix bezeichnete grafische Tool zeigt zu einem bestimmten Begriff redaktionelle Verweise

(direkte inhaltliche Verweise zu einem anderen verwandten Lexikoneintrag), ähnliche Themen (semantisch nah verwandte Themen) und entfernte Beziehungen (semantisch entfernt verwandte Themen) in einem netzwerkartig organisierten Graphen an (verschiedene Farben dienen der Unterscheidung der genannten drei Verweisarten). Dabei kann die Anzahl Knoten innerhalb des Graphen dynamisch so variiert werden, dass der Benutzer selbst entscheiden kann, wie viele Verweise (Knoten) auf einmal dargestellt werden sollen. Das Auswählen eines Knotens mit der Maus führt dann zur Einblendung eines PopUp-Fensters mit dem jeweiligen Textanfang des verwiesenen Lexikoneintrags, das Auswählen eines dort angehängten Links direkt zu dem Zieldokument. Interessant wäre natürlich die Frage, ob die erwähnten Verweise redaktionell erstellt oder ob diese durch ein automatisches Verfahren generiert werden.

3.4 Link-Verzeichnisse

ENFORUM soll über eine virtuelle Bibliothek verfügen, in die durch Anwendung der Hypertextmethodologie externe Ressourcen und Zusatzinformationen der Beitragenden eingebunden werden.

Verknüpfungen zu externen Quellen sollen in der Regel direkt bei den Glossarobjekten, den verschiedenen Typen von Wörterbucheinträgen, angebracht, also aus dem aktuellen Kontext heraus verstanden und genutzt werden. Verknüpfungen sollen die referenzierten Websites kurz beschreiben und bewerten. Darüber hinaus soll ENFORUM auch ein separates Verzeichnis der intensiv benutzten Quellen enthalten, z.B. Verweise zu anderen Glossaren/Wörterbüchern. Das Link-Verzeichnis soll automatisch auf einem aktuellen Stand gehalten werden. Kriterien des Update könnten sein: Anzahl der Verknüpfungen auf die spezielle Quelle oder Anzahl der Besuche dieser Quelle vom Wörterbuch aus. Weiterhin sollen alle Verknüpfungen regelmäßig vom System daraufhin überprüft werden, ob ihre Referenzen noch existieren. Bei Fehlermeldungen während der Benutzung sollen die Nutzer über ein Formular zur Mitteilung an das Redaktionsteam aufgefordert werden. Das Link-Verzeichnis soll durch den Einsatz von Suchrobotern, die in Online-Datenbanken und Internet-Ressourcen nach Referenzen zu schon im ENFORUM existierenden Beiträgen suchen, ständig automatisch aktualisiert werden. Die Einschlägigkeit der ermittelten Referenzen soll von den jeweiligen Autoren bzw. Moderatoren überprüft werden.

3.5 Neue Metapher für die Interaktion mit ENFORUM

Entsprechend der grundlegenden Konzeption von ENFORUM soll in Zukunft die kommunikative Komponente von ENFORUM mehr im Zentrum des Systems stehen. Wie die ersten Erfahrungen mit ENFORUM jedoch gezeigt haben, ist die Bereitschaft der Fachwelt, aktiv an der Entwicklung von ENFORUMs-Einträgen durch Begriffsvorschläge und vor allem durch Beteiligung an den Kommunikationsforen mitzuwirken, im Prinzip zwar erfreulich vorhanden – die meisten im Informationsgebiet arbeitenden Forschungs- und Ausbildungseinrichtungen haben über entsprechende Personen ihre Mitarbeit zugesagt -, faktisch lässt die aktive Beteiligung jedoch noch viele Wünsche offen. Ein elektronisches Kommunikationssystem ist, wenn es nicht wie E-mail von unmittelbarer Relevanz für den professionellen oder privaten Alltag ist, kein Selbstläufer, sondern muss Anreize bieten.

Wir versprechen uns für die Zukunft viel von dem oben erwähnten Honorierungssystem, das im wissenschaftlichen Kontext nicht auf Belohnung durch Entgelt, sondern auf Belohnung durch Reputation abzielt. Anerkennungsverfahren mit Konsequenzen des Rating der individuellen oder institutionellen Leistung werden auf Grund der sich verschärfenden Wettbewerbssituation auch im Wissenschaftssystem zum Alltag gehören, also wohl auch bei ENFORUM.

Weiterhin wird die aktive Beteiligung von dem Ausbau der oben behandelten Moderationsfunktionen abhängen. Das gilt zwar in erster Linie für die Teilnahme und dann auch aktive Beteiligung an den Kommunikationsforen, aber auch für die Gewinnung von Autoren, die durch ihre Begriffsvorschläge erst die Grundlage für eine fachliche Kommunikation schaffen. Hierzu müssen entsprechende Expertensysteme (Systeme über Experten) entwickelt sein, aus denen gezielte Hinweise auf einschlägige Experten abgeleitet werden können. Solche adaptiven Funktionen ermöglichen individuelle Ansprache und individualisierte Beiträge aus der Fachwelt.

Wir wollen hier aber auf einen weiteren Ansatz zur Verbesserung des Aktivitätsgrades in ENFORUM hinweisen. Akzeptanz - und man kann dies erweitern auf Vertrauen als Grundlage für Beteiligung - von Informations- und Kommunikationsangeboten ist auch eine Funktion der Benutzungsoberfläche (Manhartsberger/Musil2002) oder - um es auf unsere Fragestellung direkt zu beziehen - der Benutzungsmetapher, die als Prinzip der Gestaltung und damit auch als Anreiz der Kommunikation dient.



Abb. 5: Diskussionsforum zu einem Begriff

ENFORUM, abgesehen vom Navigationstool, und die meisten bisherigen elektronischen Wörterbücher im Internet verwenden die klassische Alphabetisierungsmetapher bzw. das Matching Paradigma des Information Retrieval als User Interface. D.h. der Einstieg in das System erfolgt entsprechend dem Prinzip der direkten Manipulation durch Anklicken des entsprechenden Buchstabens des grafisch aufbereiteten Alphabets oder - komfortabler und weiterführender - durch Eingabe einer Suchformulierung mit dem Ergebnis von Listen von Einträgen, in denen das oder die formulierten Suchbegriffe vorkommen.

Im Zentrum der Darstellung der Wörterbuchleistung steht die informationelle Einheit, d.h. der Begriffseintrag der jeweiligen Wörterbucheinheit. Sofern kommunikative Funktionen vorgesehen sind, werden sie meistens eher peripher behandelt, so auch bislang in ENFORUM. D.h. eine Diskussion über einen Begriffsvorschlag wird am Ende dieses Eintrags eröffnet und führt entsprechend ein gewisses Schattendasein, das nicht direkt zur aktiven Teilnahme motiviert. Zudem wird die fortlaufende Diskussion in der Regel einfach über das „Thread“-Paradigma dargestellt, d.h. ein Kommentar zu einem Ein-

trag im Kommunikationsforum wird durch Einrückung markiert, so dass schrittweise eine Hierarchie der Kommunikationsbeiträge entsteht. Orientierung über die „Subject“-Zeile ist im Prinzip möglich, wird aber durch einfache Verwendung des RE-Subject, RE-RE-Subject nicht gerade aussagefähig verwirklicht.

Wir haben vor, sozusagen eine kopernikanische Wende in der Metapher des Interaktionsverhaltens einzuleiten. Nicht länger soll mit Blick auf das Design die Informationsfunktion im Vordergrund stehen und damit die Suchfunktionen zum Einlösen des Informationsbedarfs, sondern die Kommunikationsfunktion. Ausgangspunkt zur Realisierung dieser „Wende“ könnte die Inanspruchnahme der ENFORUM schon in der Benennung zugrundeliegenden Forums-Metapher sein. Ein Forum in seinem antiken Ursprung des *forum romanum* war der Ort des Austausches der freien Bürger, die die Angelegenheiten der *res publica* zum Gegenstand ihres Diskurses hatten, natürlich auch das Forum als Marktplatz des Handels mit realen Gütern verwendeten.

Entsprechend soll auch ENFORUM die Kommunikations-/Austauschleistung in den Vordergrund stellen. Für den Benutzer von ENFORUM sollte schon im Eingang ersichtlich sein, welche Kommunikationsforen über welche Begriffsvorschläge aktuell aktiv sind. Durch direktes Aktivieren der ihn interessierenden Foren werden im Kommunikationsraum im Zentrum der Bildschirmrealisierung die aktuellen Diskussionen visualisiert, wobei das Thread-Paradigma durch ein zunächst zweidimensionales, in weiterer Zukunft räumlich dreidimensionales Netzparadigma abgelöst werden soll, auf Grund dessen frei in Kommunikationsnetzen navigiert werden kann.

Der zentrale Kommunikationsraum wird umgeben durch weitere „Räume“, die entsprechend metaphorisch gestaltet werden müssen, als da sind:

- Der Informationsraum, in dem die konsolidierten, aber auch in der Entwicklung befindlichen Begriffe/Begriffsvorschläge angezeigt werden können.
- Der interne und externe Such-/Navigationsraum, der es erlaubt, in den Beständen von ENFORUM zu suchen bzw. zu navigieren und der es darüber hinaus gestattet, in externen Informationswelten (andere Wörterbücher, externe einschlägige Fachinformationsbanken, Suchmaschinen zum Erschließen von einschlägigen Internet-Ressourcen)

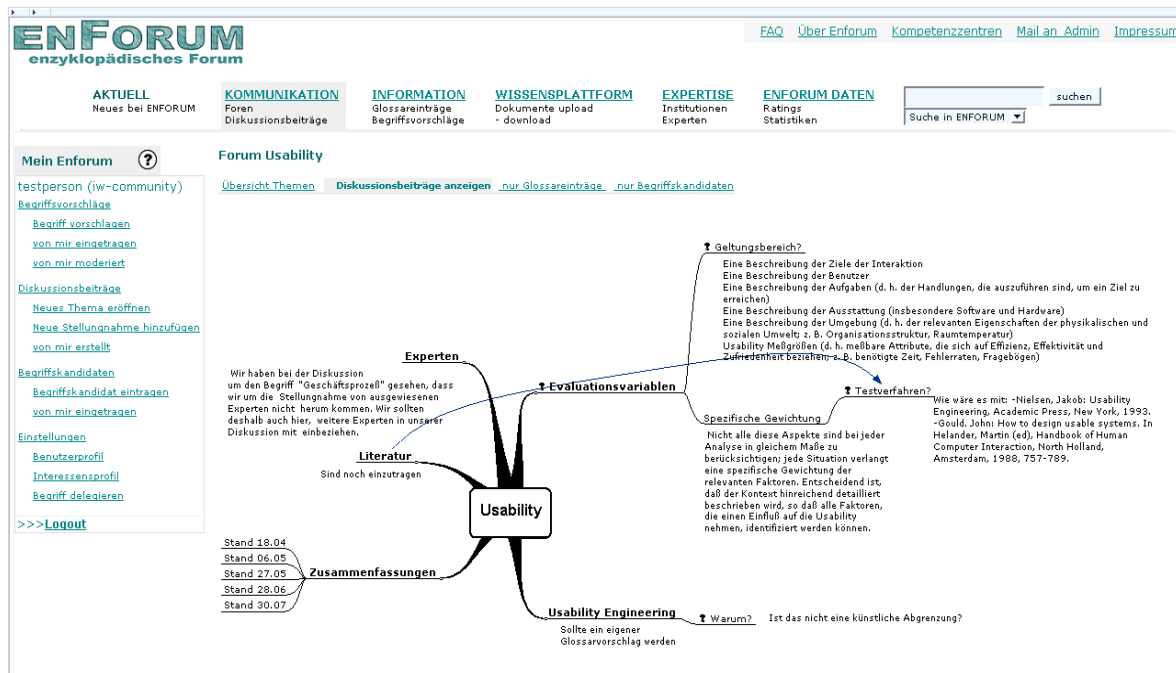


Abb. 6: Mögliche zukünftige Benutzeroberfläche von ENFORUM

- Der Expertise-Raum, durch den Personen und Institutionen, die für neue Beiträge oder für die Beteiligung an den Kommunikationsforen in Frage kommen, nachgewiesen werden können
- Der Referenzraum, der als Wissensplattform weiterführende Information zu ENFORUMs-Einträgen bereitstellt und die interaktive Möglichkeit des Herunter- und Heraufladens von Artikeln anbietet
- Der Reputationsraum, in dem darüber Buch geführt wird, auf Grund welcher ENFORUMs-Leistungen der aktuelle Rating-Stand der ENFORUMs-Teilnehmer ermittelt worden ist.
- Der Aktualitätsraum, in dem unter verschiedenen Gesichtspunkten – Eingabezeitraum, Abfragehäufigkeit, Interaktions-/Kommunikationsgrad, Informationsgrad (Anzahl der eingespielten Referenzen), ... - aktuell interessierende Beiträge angezeigt werden.

Im Rahmen des jetzigen ENFORUM könnte - nach ersten Überlegungen - eine entsprechende Benutzeroberfläche wie folgt aussehen. Eine konsequente Umsetzung in die verteilte Forum-/Raum-Metapher steht aber noch aus.

Die Reichweite und Konsequenz von neuen attraktiven Benutzungsmetaphern und deren Ausgestaltung, hier über verschiedene ENFORUMs-Räume, soll gewiss nicht überschätzt, aber auch nicht marginalisiert werden. Darf die MacLuhan'sche medientheoretische Aussage „the media is the message“ zwar nicht absolutiert, schon gar nicht deterministisch interpretiert werden, so hat sie doch einiges an Plausibilität für sich. Im Kampf um die aktive Teilnahme

an den ENFORUMs-Prozessen, die für den Erfolg des kommunikativ kooperativ verteilten ENFORUM-Vorhabens entscheidend ist, kann die innovative Kommunikations-/Forums-Metapher ein entscheidender Erfolgsfaktor sein. ENFORUM wird jedenfalls damit in der Zukunft intensiv experimentieren.

4 Literaturangaben

- [Bekavac 1999] Bekavac, B.: Suche und Orientierung im WWW (Schriften zur Informationswissenschaft, 37). Universitätsverlag Konstanz, 1999
- [Lenich 2002] Lenich, A.: Funktionen zur Orientierung in einem virtuellen, kollaborativen Wörterbuch (ENFORUM). Fachbereich Informatik & Informationswissenschaft, Universität Konstanz, 2002
- [Manhartsberger/Musil 2002] Manhartsberger, M.; Musil, S. (2002): Web usability. Das Prinzip des Vertrauens, Galileo Design: Bonn
- [Semar et al. 2002] Semar, W.; Bekavac, B.; Griesbaum, J.; Kuhlen, R.; Schütz, T.: ENFORUM – konzeptorientiertes Wissensmanagement für den Informationsbereich. In: Schmidt, Ralph (Hg.): Content in Context, Perspektiven der Informationsdienstleistung. Proceedings – 24. Online-Tagung der DGI. Frankfurt: Deutsche Gesellschaft für Informationswissenschaft und Informationspraxis, 2002. S.87-98



Personal Assistance: Helping the Users Find Their Way through the Information Space

Kurt Englmeier¹, Javier Pereira²

¹German Institute for Economic
Research (DIW)
Koenigin-Luise-Str. 5
14195 Berlin Germany

kurt@diwsysv.diw-berlin.de

² Departamento de Informática de
Gestión
Universidad de Talca
Avda. Lircay s/n
Talca, Chile
jpereira@utalca.cl

Abstract

Qualitatively enhanced access to data collections requires a more meaningful interface which endows the users with personal agents that are easily adaptable to the users' individual retrieval preferences. Tracking individual needs of users leads to personal digital assistance that appears in the system's interaction mode as the user's personal software agent. The main objective of personal assistance as envisaged here is enabling the system to give the users recommendations that are derived from best practice in searching for information. Best practice reflects the knowledge about how users perform successful retrieval activities and attempt to define the best ways of searching large document collections. Best practices are developed on an individual as well as on a community base.

The model presented here was developed for the information system IRAIA, a portal for economic information from huge data collections of Economic Research Institutes (ERI) and National Statistical Institutions (NSI). The personal agent draws on IRAIA's model of context-oriented retrieval and observe, record and analyse the users' search and navigation by applying concepts of the respective taxonomies.

1 The artefact of personal assistance

Personal assistance as discussed here is positioned within the framework of adaptable Web systems. These systems collect data for a user model from



various sources that also include implicitly observing user interaction and enabling explicit user response. [Brusilovsky & Maybury 02] Instead of focusing on a user model we more concentrate on a use model. The model entails an adaptation effect that tailors retrieval interaction to different use types. Whereas a purely adaptive system bases solely on a user or a use model the adaptable system requires the user to specify major parts of its behaviour. In the context of our application area we concentrate on adaptable navigation support. Based on a number of successful approaches in adaptive hypermedia and web systems [see Billsus et al. 02, Yang et al. 99, for instance], our design of an adaptable system is intertwined with a strong linguistic layer that raises significantly the expressiveness of the interaction mode.

Personal agency endows an information system with an easy and highly user-oriented interaction mode. This is achieved by the installation of a personal agent that helps the user during navigation. It starts with observing the user's selections of entries from the concept hierarchies, i.e. with capturing query profiles. An actual profile then is compared with already stored ones and recommendations are derived if the agent can recognise a certain similarity between profiles. The recommendation refers to further selections of concepts suitable in an actual retrieval situation. Our design of a personal agent results from combining the approaches for agents as personalised companions [André & Rist 02, Billsus & Pazzani 99] and agents for automatic text analysis [Wermter 00, for an overview see Mladenic 99] and for developing interaction strategies [Durfee 01, Englmeier et al. 01].

2 Getting orientation in complex information spaces

Information arises from data when they are combined, arranged, and presented accordingly. Only a suitable combination of related time series and texts is in the position to convey the information that is contained in these separate and otherwise imperceptible components. An interface expressing the content of these components in a comprehensive and uniform way is of outstanding importance when it comes to proliferate information that has to be composed by distributed and heterogeneous data. This holds for most of the large-scale and complex data collections in general, but for those of ERIs and NSIs in particular. Our model for the construction of context-aware information spaces is derived from related models of retrieval environments for interacting with large data collections [Agosti et al. 92, Krause 96]. Interfaces presenting such a content overview help to construct a retrieval environment where the users explore data collections within a semantic coordinate system derived from taxonomies of the respective information domain. These tax-

onomies exist for a variety of application areas. They are a solid basis for a controlled and structured vocabulary and therefore most appropriately for the content semantics defining a domain-related context.

A semantic coordinate system endows the users with a concise as well as comprehensive vocabulary. Hierarchically arranged and grouped along major content facets this vocabulary acts as a stable coordinate system easy to comprehend and memorise¹. The users are thus much more in the position to localise themselves effortlessly. Successfully searching and navigating now means guided travelling from information to information just by changing the semantic coordinates, i.e. by pointing to relevant concepts. This structure on the other hand enables to pinpoint the semantic location of any kind of information. It also supports the correct identification of retrieval strategies which extends content searching and navigating towards context-assisted retrieval.

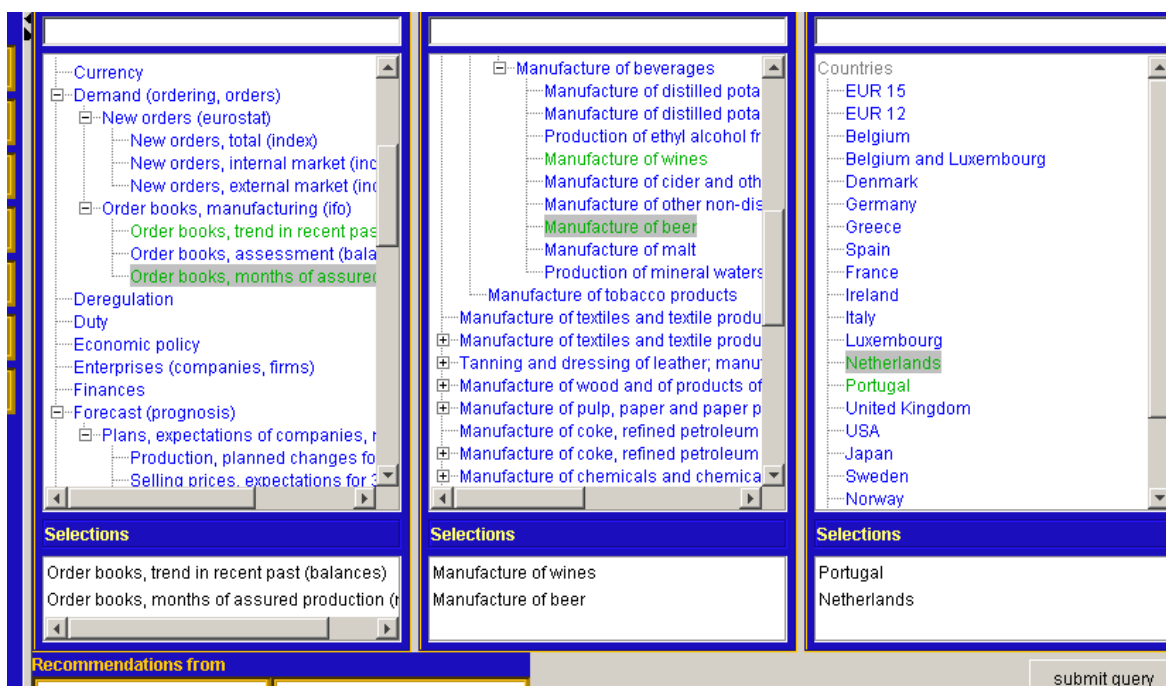


Figure 1. Searching and navigation in a semantic coordinate system. Selected concepts make up the initial query profile. While realising his retrieval strategy the user usually performs iterative steps of defining a query and analysing the retrieved results. In IRAIA, the documents are annotated solely with entries from the hierarchies. While viewing a document the interface shows also the annotated concepts (similar to the screen above) that are thus already familiar from the initial query formulation. Modifying the set of annotated concepts is thus tantamount with repetitive query formulation.

¹ The coordinate system itself can be presented simultaneously in different languages. This ensures that the domain model mentioned later features multilinguality.

Semantic coordinates can be rendered simply and efficiently by concept hierarchies displayed in adjacent windows. The user just pinpoints to the relevant concept (i.e. a phrase of terms). A group of selected concepts that is sent to the system reflects a query or query profile. In this context, a sequence of such profiles represents a navigation history. In addition to this, a particular best practice in retrieval is reflected by a set of individual navigation histories that appear quite frequently among all observed histories.

For IRAIA we produced a powerful taxonomy that merges two of the most important structures in this field: eurostat's NACE² nomenclature and the industry systematic of the IFO institute for economic research. The unified taxonomy creates a semantic coordinate system that enables exact and automatic positioning³ of coherent documents even if they are of different types. It also provides users with the necessary orientation while exploring the information space. Like in using languages it helps users as a passive vocabulary to identify the topics of their information problem.

3 Task model of the personal agent

At each decision point that is again represented by a query profile the user can ask the agent to show recommendable concepts. These concepts are part of a profile in the sequence of an archived navigation that follows the profile corresponding to the decision point. This means there is a sequence of profiles having significant concepts in common and additionally at least one profile that reaches beyond the decision point. Again, the prerequisite for being a candidate for recommendation is that such a profile has a significant overlapping navigation history with the actual sequence in the recent past of the decision point. In the figure below this profile has the number 19 and contains concepts for recommendation. This means, these concepts are marked respectively in the corresponding hierarchies presented in the user interface.

An archived profile is suitable for recommendation if it contains similar query profiles up to the corresponding decision point and if there is a query profile that reaches beyond this point. In this context, it is not important if the pro-

² Nomenclature des Activités dans la Communauté européenne -systematic of the economic activities of the European Union

³ Automatic positioning means annotating a document with entries from the concept hierarchies. During the automatic process the terms of a document are matched with those of the hierarchies and their synonyms. The hierarchy entries which got the highest frequency of matches are chosen for the annotation.

files belong to the user's personal history lists or to the common collection of best practices.

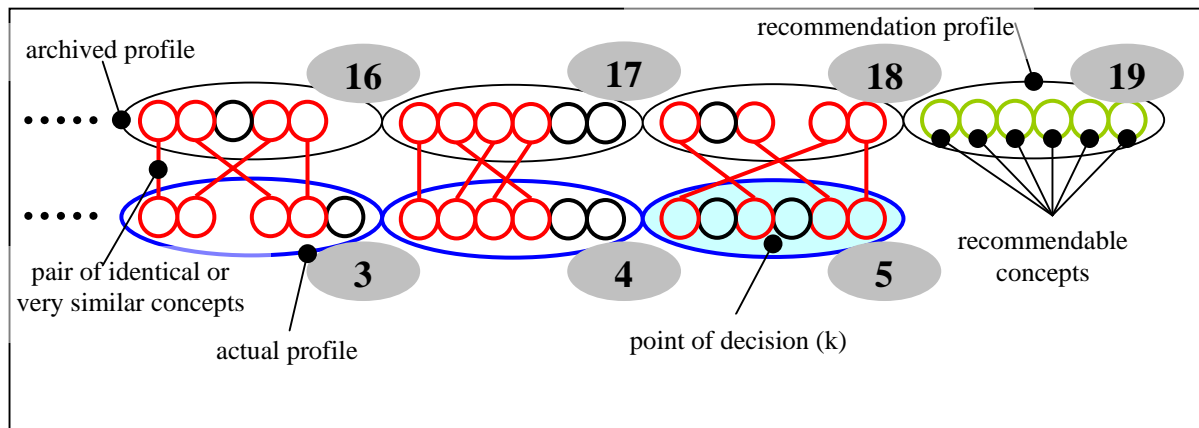


Figure 2. Schema of a recommendation situation. The upper profiles (16 through 19) are archived ones. Due to concept similarities up to the decision point (k) and the availability of a further profile (19) that reaches beyond k they are suitable for recommendation. The lower profiles reflect the actual navigation of the user. Pairs of similar concepts are marked by circles with links in between them. The order of the nodes within a query profile is not important. In this example, the users ask for recommendation while completing query no. 5 and gets concepts from profile 19.

3.1 Human-agent interaction

A profile can be regarded similar to another one if it covers at least 80% of its concepts. Taking into account that coverage varies among the profiles of an observed track the same threshold is applied to measure the overall similarity between two tracks. Among all tracks with similarity values above this threshold the one with the highest value is selected for recommendation.

The work of the agent now is

- to identify significantly similar queries,
- to decide if there is a significantly common navigation history,
- to present recommendable concepts,
- to show the user the way of their decision,
- to adapt its way of deciding by asking the user whether the recommendation was useful, potentially useful or not useful, and
- to accept new decision rules.

It is important to note that the archived sequences may reflect the user's individual navigation history or best practices from the user community. Tasks and decisions outlined above are the same for both cases even if different kinds of agents are in charge with. During navigation, recommendable concepts are marked respectively.



Figure 3. The user can ask the personal agent to look for recommendations in its local archive of personal navigation data or to ask remote agent for a suggestion derived from best practices.



Figure 4. Nodes are marked according to the concepts of recommendation profile.

3.2 A decision model for the personal agent

Architecture of the agent system

The multi-agent environment in IRAIA addresses agents interacting in order to produce suggestions to the users navigating information spaces that are presented by theme specific taxonomies.

The evaluation criteria of recommendations

Let us define $O = \{o_1, o_2, \dots, o_n\}$ as the user's (observed) navigation track and $T_j = \{t_{j1}, t_{j2}, \dots, t_{jm_j}, \dots, t_{jn_j}\}$ ⁴ as a recommendation from the archived sequence j of all available best practices or private navigation histories. Remember that track elements correspond to profiles containing nodes of the concept hierarchies. At each navigation step – i.e. at each query – the users select a number of concepts that constitutes a query profile.

An important role has the recommendation profile, when after a number of navigation steps the users need support for the decision which concepts to choose next. At this decision point k the user ask for recommendations that should be based on representative and helpful navigation examples from their own or the community's navigation histories.

We assume that a (community or individual) navigation history T_j is adequate for a user's retrieval if it contains in its first part (up to the point m_j that corresponds to the decision point k) as many profiles as possible similar to those of

⁴ o_k and t_{jm_i} are query profiles; O and T_j contain n and n_j profiles respectively.

O and, at the same time, only a few different ones. Additionally, if there is a profile beyond t_{jm_j} , we may consider it as interesting to the user. Thus a recommendation for further navigation comprises all the nodes of the first profile of T_j beyond the point t_{jm_j} , that is t_{jm_j+1} . In general, it can be assumed that as closer the profile comes to the decision point the more useful it is for recommendation.

We assume that an information problem triggering and steering the user's navigation can be divided into a smaller number of sub-problems that can be observed frequently over a number of more complex problems. Even if two users have the same information problem they satisfy their need in a different way resulting in different sequences of profiles. Certain clusters among these sequences, however, are very similar. We therefore apply our analysis to these "atomic" problems. At the same time it can be assumed that the user needs decision support that addresses the most recent navigation history that usually coincides more with such an "atomic" problem rather than a larger sequence of past navigation steps. Let $O_c = \{o_{k-L}, \dots, o_{k-1}, o_k\} \in O$ be a sequence of profiles in the actual navigation and $T_{jc} = \{t_{jm_j-L}, \dots, t_{jm_j-1}, t_{jm_j}\} \in T_j$ the part to be compared from an archived navigation (of best practices or personal history).

Then the agent resorts to a certain archived track if it complies the following conditions⁵.

1. Sufficient similarity between pairs of profiles.

Let us assume that I_{O_c} and $I_{T_{jc}}$ are the indexes of concepts in O_c and T_{jc} respectively, w_i the weight of a concept i , C the set of indexes of O_c also contained in T_{jc} , and D the set of indexes of O_c not contained in T_{jc} .

Then, we define the rate of common concepts as $\phi = \frac{\sum_{i \in C} w_i}{\sum_{I_{O_c}} w_i}$.

In the same way we define the rate of distinct concepts as $\delta = \frac{\sum_{i \in D} w_i}{\sum_{I_{T_{jc}}} w_i}$.

Therefore, a sufficient similarity is expressed here by $\phi \geq 0.5$.

⁵ All the values mentioned here are just initial ones. They all have to be approved by later tests.

If two profiles $t_v \in T_{jc}$ and $o_u \in O_c$ fulfil this criteria, we say both are similar or $t_v \approx o_u$.

Thus, we can define the following similarity function $\sigma = \phi(1 - \delta)$

If O_c and T_{jc} contain exactly the same concepts holds $\sigma = 1$, whereas it holds $\sigma = 0$ if they are completely different. According to our model we may put $|O_c| = |T_{jc}|$, $w_i = \frac{1}{|O_c|} \forall i \in I_o$ and additionally the half of concepts are common, then we have $\sigma = 1/4$. In other words, the existence of distinct concepts impose a penalty to the similarity aggregated function. The more the distinct concepts in T_{jc} are important, the bigger this penalty.

These criteria are evaluated independently of profiles, that is, we use all concepts in O_c and T_{jc} . It can be clearly a problem, but it permits to do a first evaluation of tracks T_{jc} .

Thus, the decision processes of the agents are single filtering processes. In such case, we think the user have no deal with this filtering and they must remain ignorant about the agent's logic. Only the experts may set the values of these parameters.

2. Sufficient similar history among both sequences of profiles O_c and T_{jc} .

O_c and T_{jc} are similar if:

$\forall o_u \in O_c \wedge \forall t_v \in T_{jc} : o_u \approx t_v$ with $k - L \leq u \leq k$ and $m_j - L \leq v \leq m_j$. For the time being we put $L \geq 2$.

3. The **ordering** of concepts within a profile o_u or t_v , as well as the ordering of profiles within O_c or T_{jc} is irrelevant.

4. Availability of recommendation profiles.

$o_k \approx t_{jm_j}$ and $\exists t_{jm_j+1}$. This means that the decision point may lay anywhere in O as long as there is a corresponding point m_j in T_{jc} that has a profile in $m_j + 1$.

Let l_{uv}^j be the shortest path from $o_u \in O_c$ to $t_v \in T_{jc}$ ($t_v \notin O_c$). Then, total distance to non common concepts is expressed by $\theta = \sum_v \min_u \{l_{uv}^j\}$. Thus we have three criteria used for evaluate the recommendation quality: $g_1(T_{jc}) = \phi$, $g_2(T_{jc}) = \delta$ and $g_3(T_{jc}) = \theta$.

Note that the similarity and dissimilarity functions permit to establish what a practices are recommendable or not:

- $g_1(T_{jc}) \geq \xi_1$ and $g_2(T_{jc}) < \xi_2$, T_{jc} is a recommendable practice.
- $g_1(T_{jc}) \geq \xi_1$ and $g_2(T_{jc}) \geq \xi_2$, T_{jc} is not a recommendable practice.
- $g_1(T_{jc}) < \xi_1$ and $g_2(T_{jc}) < \xi_2$, T_{jc} is not a recommendable practice.
- $g_1(T_{jc}) < \xi_1$ and $g_2(T_{jc}) \geq \xi_2$, T_{jc} is not a recommendable practice.

Here, $\xi_1 \in [0.5, 1)$, $\xi_2 \in (0, 0.5]$ are thresholds representing the expected similarity and dissimilarity bounds, respectively, for the recommended practices (or archived sequences).

Elaborating suggestions to user

Evaluating and selecting recommendations for the user means assigning best practices selected by the personal agent according to one of three classes: interesting (C_3), potentially interesting (C_2) or non interesting (C_1) suggestions. A category is explained by one (or two) bounder profile, i.e. a vector where each component indicates the evaluation on a criterion for a particular recommendation situation (see Fig. 7). Let b_h ($h = 1, 2$) a category bounder profile; then, it is perfectly defined, in terms of criteria, by the vector $(g_1(b_h), g_2(b_h), g_3(b_h))$.

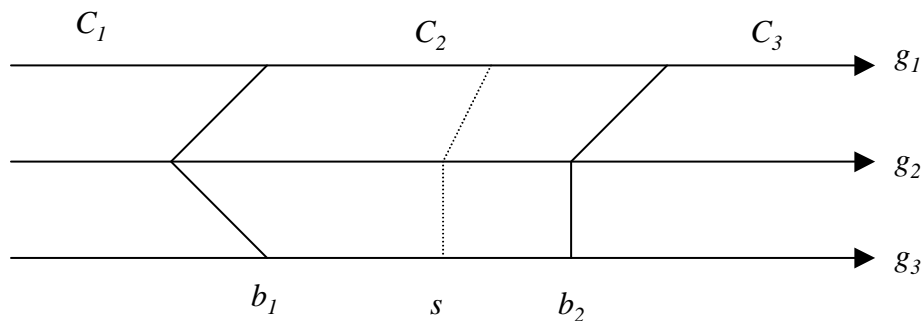


Figure 7. Criteria, categories and bounders of category profiles

Thus, let $s \in S$ be a suggested best practice and S a binary preference relation, nominated an *outranking relation*, such that the expression $s \geq b_h$ means that s is “at least as good as” b_h , i.e. $(g_1(s), g_2(s), g_3(s))$ is globally evaluated, in terms of preferences, as similar or right-side from $(g_1(b_h), g_2(b_h), g_3(b_h))$. For instance, we have $s \geq b_1$, but $not(s \geq b_2)$.

At this point, we propose to use the pessimistic ELECTRE TRI method [Mousseau et al. 01] to assign best practices. This method bases on a paired comparison between suggested best practices and the category bounder pro-

files. Therefore, in the pessimistic assignment procedure $s \in S$ is assigned to C_{h+1} if b_h ($h = 2,1$) is the first profile where $s \geq b_h$. Actually, in ELECTRE TRI, one must set an index $\sigma(s, b_h)$ reflecting the credibility of statement $s \geq b_h$. A cutting level λ is also set to found the minimal credibility which permits to validate the statement. The next relations describe the global comparison among a suggested best practice and a bounder profile:

$$\begin{aligned} \sigma(s, b_h) \geq \lambda \text{ and } \sigma(b_h, s) \geq \lambda &\Rightarrow s \geq b_h \text{ and } b_h \geq s ; s \text{ assigned to } C_{h+1} \\ \sigma(s, b_h) \geq \lambda \text{ and } \sigma(b_h, s) < \lambda &\Rightarrow s \geq b_h \text{ and } \text{not}(b_h \geq s) ; s \text{ assigned to } C_{h+1} \\ \sigma(s, b_h) < \lambda \text{ and } \sigma(b_h, s) \geq \lambda &\Rightarrow \text{not}(s \geq b_h) \text{ and } b_h \geq s ; s \text{ not assigned to } C_{h+1} \\ \sigma(s, b_h) < \lambda \text{ and } \sigma(b_h, s) < \lambda &\Rightarrow \text{not}(s \geq b_h) \text{ and } \text{not}(b_h \geq s) ; s \text{ not assigned to } C_{h+1}. \end{aligned}$$

4 Conclusion

In this paper, we presented a model of user involvement that goes beyond the scope of the methods of usability engineering. It means placing the users in the control loop by enabling them to manipulate retrieval protocols that otherwise would be compiled and applied automatically. This outstanding trait requires the use of personal assistance that is made possible by adaptable software agents. These personal software companions are available to the users just-in-time, on-demand and help users, for instance, with refinement of retrieval strategies. The assistance is derived from best practice with individual or community-wide information search strategies without losing the view on privacy issues. Software companions facilitating this assistance are adaptable to the users' individual retrieval needs and preferences. This kind of personalisation will be more and more an issue for future emerging systems.

In addition, user involvement as envisaged here avoids having to design systems in which retrieval tasks are completely automated; some processes already observed are extremely difficult to automate. Although personalisation alleviates this, the result is an ad hoc compilation of the user's process descriptions.

We tested the outlined implementation of our model within the user evaluation phase of IRAIA, majorly to discuss privacy concerns of personalisation. Threshold values as presented here emerged from the estimation of the experts involved in the project. The assistance functionality was welcomed, in general. However, there were always privacy concerns raised that focused in our case on business critical information not being treated confidentially enough. This problem could be mitigated by realising personal assistance exclusively through locally installed software components with local storage of

personal data. For the majority of users the current browser-based systems do not provide sufficiently for confidentiality.

5 References

- [Agosti et al. 92]. Agosti, M.; Gradegnio, G.; Marchetti, P. "A hypertext environment for interacting with large databases." *Information processing and management*, 28 (1992), 371-387.
- [André & Rist 02]. André, E.; Rist, T. "From adaptive hypertext to personalized web companions." *Communications of the ACM*, 45/5 (2002), 43-46.
- [Billsus et al. 02]. Billsus, D.; Brunk, C.; Evans, C.; Gladish, B.; Pazzani, M. "Adaptive interfaces for ubiquitous web access." *Communications of the ACM*, 45/5 (2002), 34-38.
- [Billsus & Pazzani 99]. Billsus, D.; Pazzani, M. "A personal news agent that talks, learns and explains." In: Etzioni, O. et al (eds), *Proceedings of the third annual conference on autonomous agents*. Seattle (USA). 1999. 268-275.
- [Brusilovsky & Maybury 02]. Brusilovsky, P.; Maybury, M. "From adaptive hypermedia to the adaptive web." *Communications of the ACM*, 45/5 (2002), 31-33.
- [Durfee 01]. Durfee, E. "Scaling up agent coordination strategies." *Computer*. 34/7 (2001) 39-45.
- [Englmeier et al. 01]. Englmeier, K.; Mothe, J.; Pauer, B. "Users bootstrap searching the Web through interactive agents supporting best practice sharing." In: M.J. Smith; G. Salvendy; D. Harris; R.J. Koubek. *Usability Evaluation and Interface Design: Cognitive Engineering, Intelligent Agents and Virtual Reality*. Mahwah, USA (2001). 923-927.
- [Krause 96]. Krause, J. "Informationserschließung und -bereitstellung zwischen Deregulation, Kommerzialisierung und weltweiter Vernetzung. ("Schalenmodell")." *IZ-Arbeitsbericht Nr. 6*. 19 (1996).
- [Mladenec 99]. Mladenec, D. "Text-learning and related intelligent agents: a survey." *IEEE INTELLIGENT SYSTEMS*, 14/4 (1999), 44-54.
- [Mousseau et al. 01]. Mousseau, V.; Figueira, J.; Naux, J. "Using assignment examples to infer weights for ELECTRE TRI method: some experimental results." *European Journal of Operational Research* 130 (2001) 263-275.
- [Wermter 00]. Wermter, S. "Neural network agents for learning semantic text classification." *Information Retrieval*, 3/2 (2000), 87-103
- [Yang et al. 99]. Yang, Y.; Carbonell, J.; Brown, R.; Pierce, T.; Archibald, B.; Liu, X. "Learning Approaches for Detecting and Tracking News Events." *IEEE INTELLIGENT SYSTEMS*, 14/4 (1999), 32-43.

Acknowledgement

Research outlined in this paper is part of the projects eStage and IRAIA that were supported by the European Commission under the Fifth Framework Programme (IST-2000-28314 and IST-1999-10602). However views expressed herein are ours and do not necessarily correspond to the eStage or the IRAIA consortium.



Virtuelle Universität Regensburg Version 2 - Internet-Plattform für Learning on Demand

Klaus J. Schäfer

Universität Regensburg
Lehrstuhl für Wirtschaftsinformatik III
Universitätsstraße 31
D-93053 Regensburg
Tel.: +49 941 943-3208
ks@wiwi.uni-regensburg.de

Zusammenfassung

E-Learning-Plattformen und -Portale sind mittlerweile auch im deutschsprachigen Raum keine Seltenheit mehr. Die Zielsetzung der hier vorgestellten *Virtuellen Universität Regensburg Version 2* (VUR 2) liegt jedoch nicht in einer ausschließlich virtuellen Lehre, sondern in der Ergänzung und Vertiefung des bestehenden Unterrichts. Hierzu kann der Inhalt von den Dozenten der Präsenz-Hochschule selbst erstellt und ohne spezielles Computer-Wissen durch einfache Webformulare im Internet bereitgestellt werden. Auf Studentenseite steht neben allen wichtigen synchronen und asynchronen Kommunikationsdiensten des Internets ein Framework für Learning on Demand mit Video- und Audiounterstützung zur Verfügung. In diesem Beitrag wird die Architektur, Funktionalität und Technik der Internet-Plattform VUR 2 vorgestellt.

Schlüsselworte: Virtuelle Lehre, Virtuelle Hochschule, E-Learning Plattform, Learning on Demand.

1 Einleitung

In den Medien werden derzeit viele Wortkombinationen mit einem vorangestellten „E“ für Electronic oder „M“ für Mobile verwendet. So z.B. „E-Business“ oder „M-Commerce“. Diese sog. „Buzz-Words“ stehen für eine Entwicklung im Bereich des Internets und der damit in Verbindung stehenden Funknetze. Wenn man den Autoren dieser Artikel Glauben schenken darf, dann liegen in diesen Bereichen wichtige Geschäftsfelder, die es in der Zukunft zu erschließen gilt. Auch der Bildungssektor kann und darf sich vor die-



ser Entwicklung nicht verschließen. So sind im Bereich des „E-Learning“ auch in Deutschland, mit etwas Verzögerung zum angloamerikanischen Raum, virtuelle Lernplattformen und Lernmanagement-Systeme auf dem Vormarsch. Mehrere Hochschulen und Universitäten sowie Verbünde (z.T. unter Beteiligung der Wirtschaft) engagieren sich im deutschsprachigen Raum sowohl mit Eigenentwicklungen, als auch mit aus den USA stammenden Plattformen.

Auch am Lehrstuhl für Wirtschaftsinformatik III der Universität Regensburg entschloss man sich 1998 nach bereits 3-jähriger Erfahrung im Bereich Tele-teaching aufgrund der mangelnden Verfügbarkeit eines passenden Produkts zu dieser Zeit, eine Eigenentwicklung vorzunehmen. Diese E-Learning-Plattform, die im Rahmen einer Dissertation entstand, wurde im Oktober 1999 als Virtuelle Universität Regensburg (VUR) erstmals vorgestellt. Version 1 wurde am 30. Oktober 1999 im Netz zur Benutzung freigegeben und kann somit als eines der ersten virtuellen Lern-Portale bzw. Lernplattformen im deutschsprachigen Raum bezeichnet werden. Seit Anfang Oktober 2001 liegt nun die Version 2 des Portals vor.

Im folgenden Kapitel wird zunächst kurz das Konzept der VUR 2 vorgestellt. Kapitel 3 gibt dann einen Überblick über die Funktionen der Plattform und beschreibt anhand von Hypertext-Flowcharts die Strukturkomponenten der Plattform. Kapitel 4 geht auf die technische Umsetzung des Systems ein. Resümee und Ausblick runden schließlich den Beitrag ab.

2 Konzept und Einordnung der Lernplattform

2.1 Konzept

Die Lernplattform VUR 2 ist nicht auf das reine Online-Studium ausgerichtet. Der Fokus liegt vielmehr auf der Ergänzung und Unterstützung der Präsenzlehre, die somit um eine hybride Lernweise erweitert wird. Hintergrund dieser Ausrichtung sind langjährige Erfahrungen des Lehrstuhls Wi III der Universität Regensburg im Bereich des Teleteaching bzw. Telelearning [LEHN01, LEHN02 u. SCHA00]. Ein Online-Studiengang mag durchaus sinnvoll sein und kann für berufsbegleitendes Lernen eine wichtige Alternative darstellen, vor allem unter dem wichtigen Stichpunkt „Lebenslanges Lernen“. Jedoch haben Studien gezeigt, dass der Präsenzunterricht, sofern die Wahlmöglichkeit besteht, von einem Großteil der Studenten bevorzugt wird. Trotz dieser Ergebnisse hat die virtuelle Lehre (und damit das hier beschriebene Portal) ihre Existenzberechtigung. Sie bietet orts- und zeitunabhängiges Lernen, das

für viele Studenten aufgrund der allgemeinen Lebenssituation (z.B. Finanzierung des Studiums bzw. Lebensunterhaltes durch Nebenjobs) geradezu einen Schlüssel zum Studium darstellt und ihnen Unterstützung bietet, das Studium in der Regelstudienzeit abzuschließen. Aber auch für Behinderte werden dadurch neue Wege eröffnet, ein reguläres Studium absolvieren zu können.

Das Potential der vorliegenden Plattform liegt jedoch in der Ergänzung der Präsenz-Veranstaltungen des jeweiligen Dozenten. Ihm obliegt damit auch die didaktische Verantwortung, was einen ganz wesentlichen Unterschied zu anderen Lernplattformen darstellt, die ausschließlich auf Fernlehre setzen und damit diese Aufgabe systemseitig unterstützen müssen. Ein weiterer Unterschied zu anderen Lernplattformen liegt in der Einfachheit der Benutzung. Da das System bedingt durch seine Ausrichtung von allen Dozenten administriert wird (und nicht nur von einem Spezialisten-Team), ist es notwendig hierfür ein sehr einfaches und übersichtliches Interface zur Verfügung zu stellen, das ohne spezielle Schulung von jedem benutzt werden kann. Um dies zu erreichen ist es notwendig, den Funktionsumfang auf ein überschaubares Maß zu begrenzen. Die anderen auf dem Markt befindlichen Systeme bieten hier meist eine Vielzahl von Funktionen, die das Interface schnell unübersichtlich erscheinen lassen, jedoch im vorliegenden Anwendungsfall gar nicht benötigt werden. Auf der Seite der Studenten gestaltet sich dies ähnlich. Auch hier ist ein einfach zu bedienendes und zusätzlich personalisierbares Interface vorhanden, über das alle Lehrmittel in im Internet gängigen Dateiformaten zum Download bzw. zur Online-Nutzung angeboten werden (so z.B. PDF, REALVideo, REALAudio, etc.).

Ein weiterer Punkt zur Abgrenzung des Systems ist der Kostenfaktor. Das System selbst ist unter der Copyleft-Lizenz (GPL Version 2) als OpenSource verfügbar und damit frei erhältlich. Die verwendete Serversoftware ist größtenteils ebenfalls OpenSource oder unter einer Educational-Licence für Bildungseinrichtungen kostenfrei. (siehe hierzu auch Kapitel 0)

2.2 Einordnung

Aus den unterschiedlichen Konzepten der vorhandenen virtuellen Lernumgebungen haben sich mittlerweile einige idealtypische Strukturen herausgebildet. Diese können auch als institutionelle Sicht bezeichnet werden. Abbildung 1 zeigt vier verschiedene Erscheinungsformen von virtuellen Hochschulen. Sie versuchen eine Einteilung aus institutioneller Sicht, wobei sich die verschiedenen Typen nicht gegenseitig ausschließen. Es sind auch Mischformen möglich, bei denen eine Hochschule mehreren Typen zugeordnet sein kann.

Aus institutioneller Sicht lässt sich damit die vorliegende Plattform VUR 2 als Typ zwei in der vorgestellten Systematik einordnen.

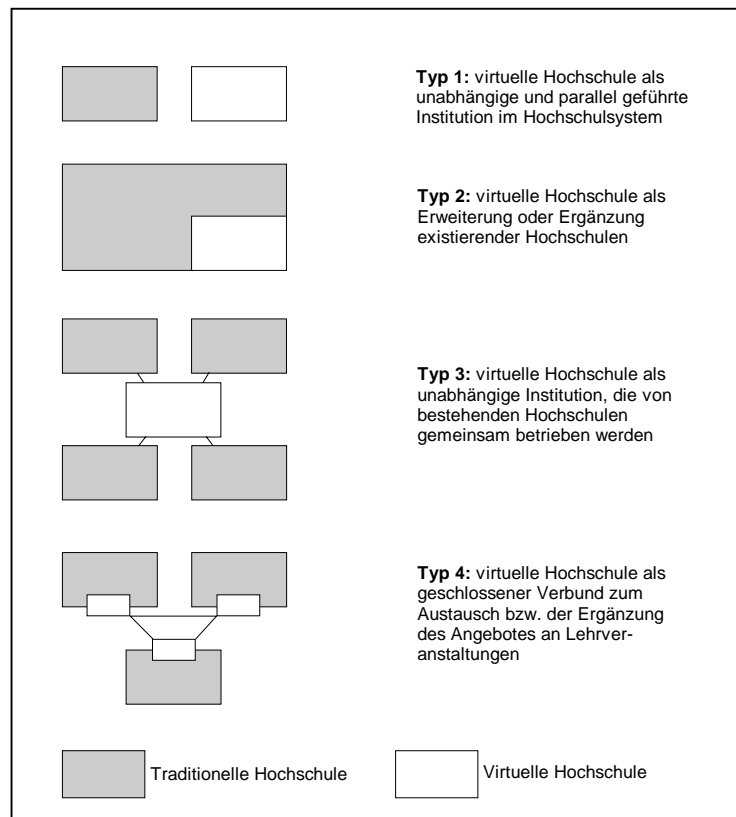


Abbildung 1: Erscheinungsformen von virtuellen Hochschulen [LEHN00]

3 Funktionen der Lernplattform

Zur detaillierten Beschreibung der VUR werden die einzelnen Funktionsbäume der Bereiche vorgestellt (eine ausführlichere Dokumentation der Funktionen ist bei [SCHA01, S. 53-56 u. SCHA02, S. 3-5] zu finden). Im Anschluss daran wird die Struktur der Systemkomponenten dargestellt.

3.1 Funktionsbäume

Die Abbildungen 2 - 4 zeigen die Funktionsbäume der VUR 2. Der Hauptfunktionsbaum (Abbildung 2) besteht aus dem Studentenbereich, der die Funktion „Teleteaching“ repräsentiert, und dem Dozentenbereich, der die Funktion „Lehrangebot verwalten“ darstellt.

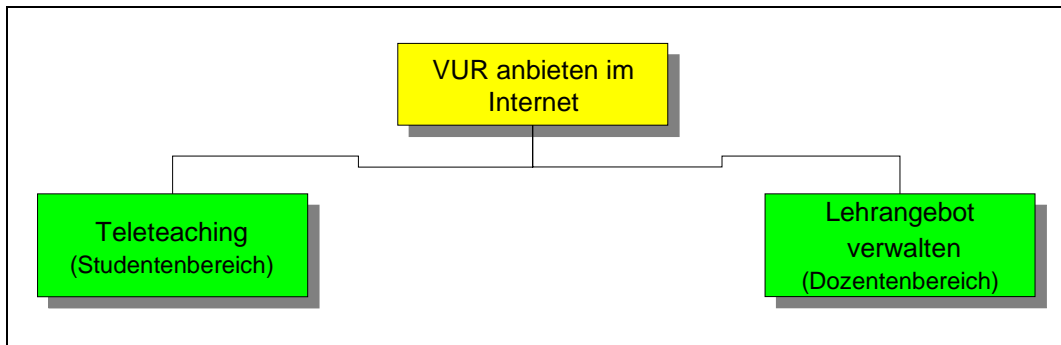


Abbildung 2: Hauptfunktionsbaum

In Abbildung 4 sind die Funktionen des Studentenbereichs zu sehen, der die Funktion „Teleteaching“ repräsentiert.

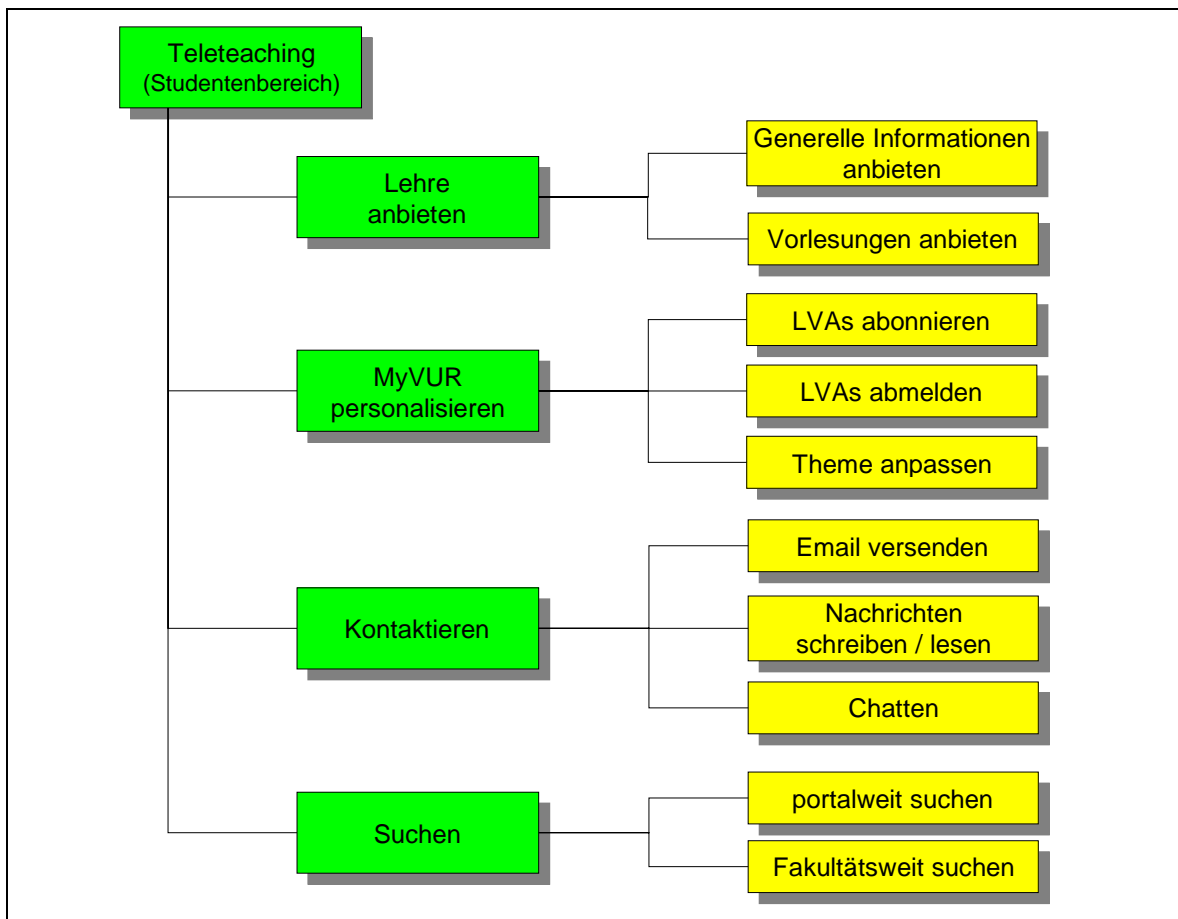


Abbildung 3: Funktionsbaum des Studentenbereichs (Teleteaching)

Die Funktion „Teleteaching“ des Studentenbereichs gliedert sich in vier Unterfunktionen, „Lehre anbieten“, „MyVUR personalisieren“, „Kontaktieren“ und „Suchen“, die wiederum in die genannten Unterfunktionen aufgespalten sind. Abbildung 5 auf der nächsten Seite zeigt den Funktionsbaum des Dozentenbereichs der die Funktion „Lehrangebot verwalten“ darstellt.

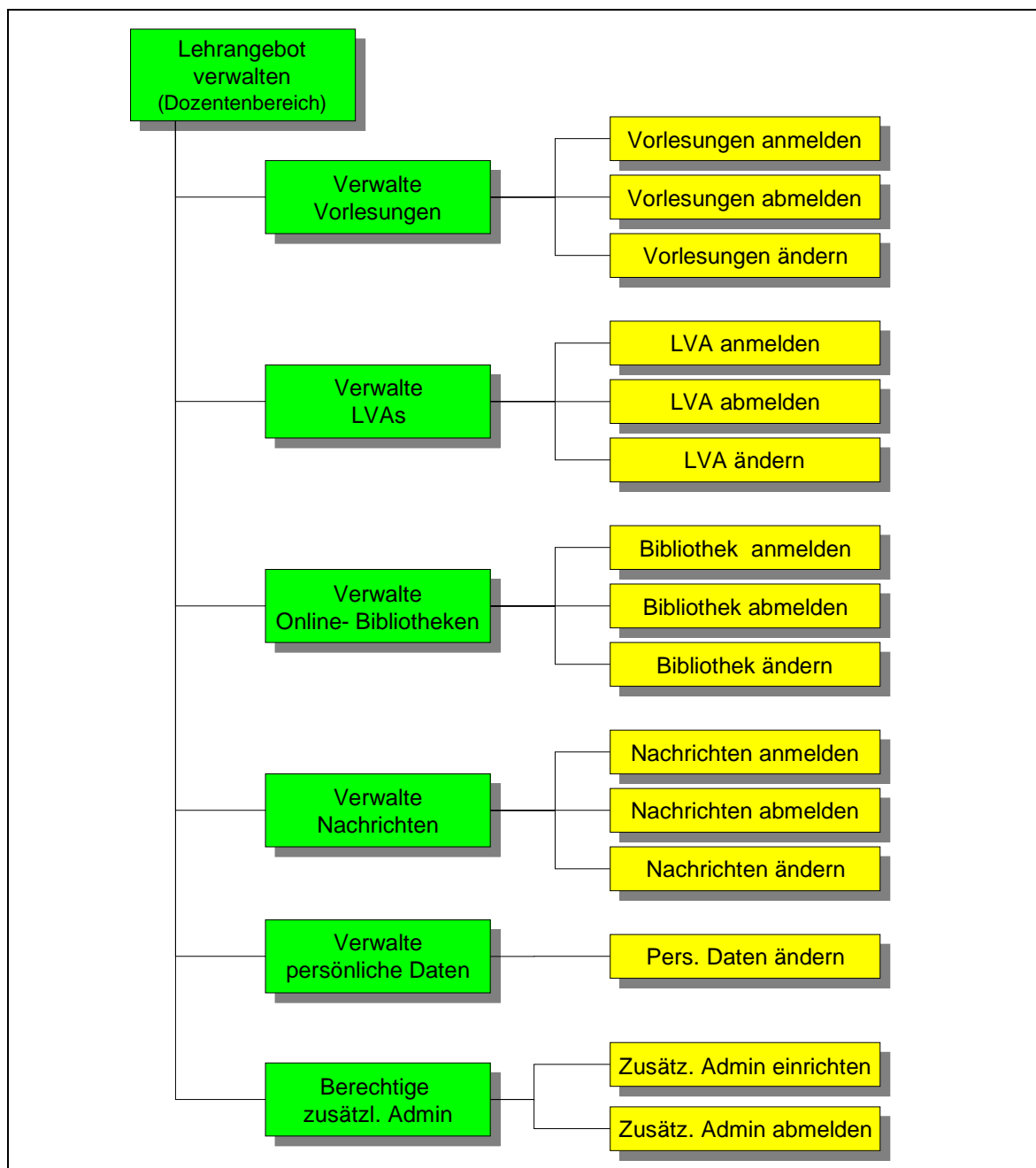


Abbildung 4: Funktionsbaum des Dozentenbereichs

Die Funktion „Lehrangebot verwalten“ gliedert sich in sechs weitere Unterfunktionen. „Verwalte Vorlesungen“, „Verwalte Lehrveranstaltungen“, „Verwalte Online-Bibliotheken“ und „Verwalte Nachrichten“ sind dabei in jeweils drei gleichlautende Unterfunktionen aufgegliedert: Anmelden, Abmelden und Ändern.

3.2 Struktur der Systemkomponenten

In diesem Kapitel soll die Struktur der Systemkomponenten anhand der Hypertext-Flowcharts des Portals beschrieben werden (Abbildung 5), wobei der Dozentenbereich nur eingezeichnet, aber nicht ausmodelliert ist. Dieser wird dann in Abbildung 6 dargestellt.

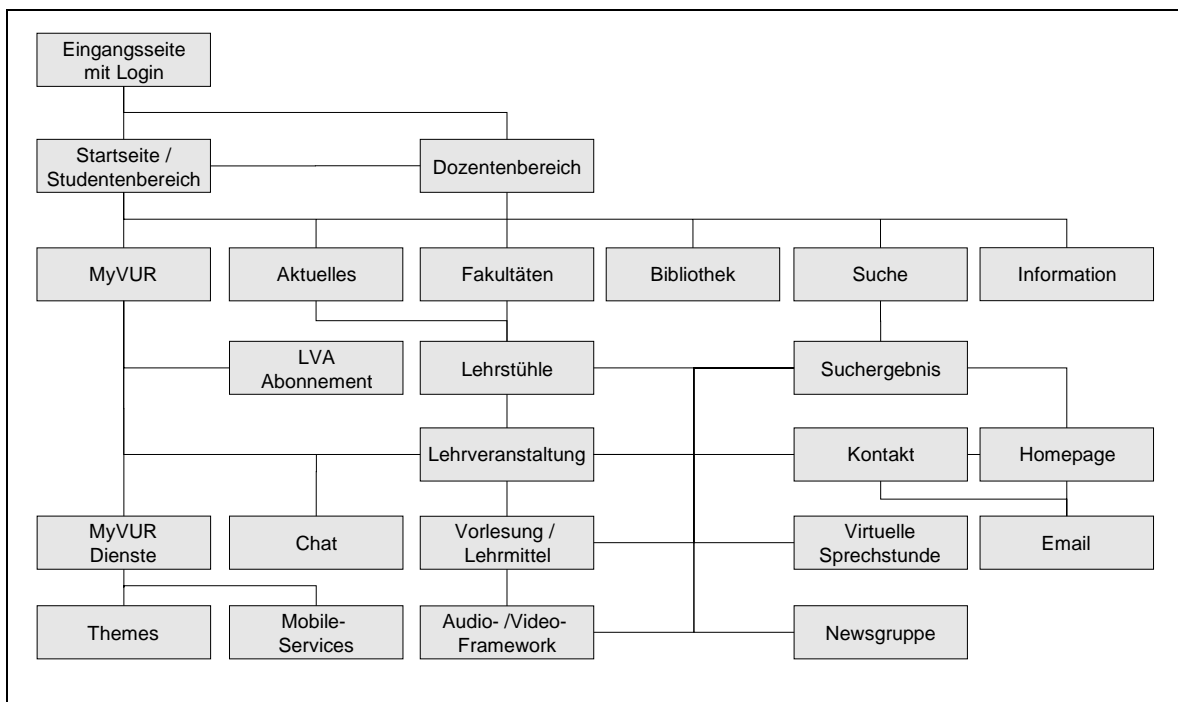


Abbildung 5: Hypertext-Flowchart der VUR 2 (ohne Dozentenbereich)

Die Eingangsseite der VUR 2 stellt neben einem hier nicht beschriebenen Gastzugang, der funktional dem Studentenbereich gleicht, zwei Zugänge bereit. Zum einen den Standard-Login für alle Studenten, zum anderen der Dozenten-Login für die Lehrbeauftragten, die zusätzlich im Dozentenbereich ihre Lehrmaterialien administrieren können. Der Dozentenbereich ist in der Hypertext-Flowchart in Abbildung 6 zu sehen. Von der Startseite, wie auch von jeder anderen Seite des Portals aus, sind die Bereiche MyVUR, Aktuelles, Fakultäten, Bibliotheken, Suche und Information direkt anwählbar. Der personalisierte Bereich MyVUR bietet die Möglichkeit der Anpassung der VUR an die Bedürfnisse des Studenten. Neben einer farblichen Anpassung durch sog. Themes, ist das Abonnement besuchter Vorlesungen für den direkten Zugriff mit „einem Klick“ möglich. Ebenso können in diesem Bereich die Mobile-Services des Portals eingerichtet werden. Der Bereich Aktuelles ist gegliedert in einen allgemeinen und in einen Lehrstuhl spezifischen Teil. Der Bereich Fakultäten listet alle verfügbaren Fakultäten auf. Von dort aus kann auf die registrierten Lehrstühle der jeweiligen Fakultät zugegriffen werden.

Diese sind weiter untergliedert in Lehrveranstaltungen (LVAs). Von der Lehrveranstaltungsseite aus ist ein direkter Zugriff auf die jeweilige Kontaktseite des Dozenten bzw. des ganzen Lehrstuhls möglich. Dort stehen dann die bereits angesprochenen synchronen (Virtuelle Sprechstunde auf Basis eines Chatsystems) und asynchronen (Newsgruppe und Email) Kommunikationsdienste zur Verfügung. Die LVA-Seite bietet eine Liste aller verfügbaren Vorlesungen bzw. Lehrmittel an. Ebenso ist von dieser Seite aus das Audio- bzw. Video-Framework startbar, in dem virtuelle Vorlesungen im Web dargestellt werden können. Das Framework bietet auf Basis der Streaming-Technologie von REAL Networks Audio- und Video-Streams in einer der Anbindung des Client-Computer entsprechenden Qualität an. Daneben sind die Folien des Dozenten sichtbar.

Die Suchfunktion bietet neben einer portalweiten auch eine eingeschränkte Suchfunktion innerhalb der registrierten Fakultäten an. Das Suchergebnis spiegelt nicht nur einen Link zur entsprechenden Ressource wieder, sondern ebenfalls alle auf dieser Seite angebotenen Links zu weiteren Diensten oder Funktionen. Der Bereich Bibliothek bietet eine vom Dozenten administrierbare Linkliste zu Online-Bibliotheken und dem OPAC-Bibliotheks-System der Universitäten an. Abbildung 6 zeigt das Hypertext-Flowchart des Dozentenbereichs.

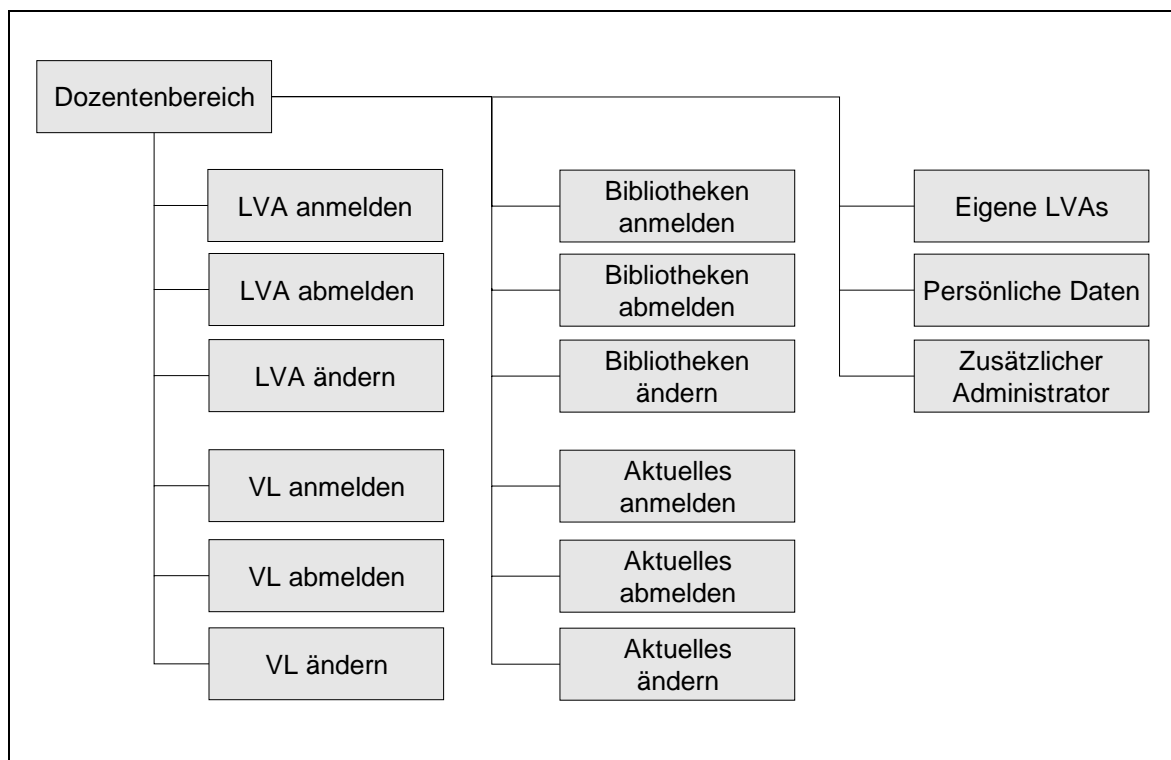


Abbildung 6: Hypertext-Flowchart des Dozentenbereichs der VUR 2

Der Dozentenbereich ist für die Lehrbeauftragten nur nach vorheriger Anmeldung und Verifizierung durch den Supervisor erreichbar (ebenfalls im Web realisiert). In diesem Bereich stehen dann die unterschiedlichen sog. „Wizards“ auf Basis von Webformularen zur Verfügung, um das elektronische Lehrangebot zu administrieren. Es können vollständige Lehrveranstaltungen mit Newsgruppen (sowohl über das Webinterface als auch über herkömmliche Newsreader zu erreichen) und Informationstexten angelegt werden. Ebenso können LVAs angepasst oder gelöscht werden. Die Vorlesungen bzw. Lehrmittel können durch einfaches Ausfüllen von Webformularen im System angemeldet, geändert oder abgemeldet werden. Dies betrifft nicht nur herkömmliche elektronische Lehrmittel, wie z.B. Skripten in unterschiedlichen Dateiformaten, sondern auch vollständige virtuelle Vorlesungen mit Video bzw. Audio-Streams inklusive der zugehörigen Foliensätze. Die Bereitstellung von aktuellen Nachrichten bzw. der Links zu den Online-Bibliotheken und deren Administration lässt sich nach dem gleichen Prinzip durchführen.

Zusätzlich besteht für den Dozenten die Möglichkeit sich alle Studenten mit Name und Email-Adresse anzeigen zu lassen, die seine Lehrveranstaltung im System abonniert haben. Die zu dieser Darstellung benötigten Informationen (Name, Matrikelnummer und Adresse) werden über den LDAP-Server der Universität bezogen.

Neben diesen vielfältigen Anpassungs- und Administrationsmöglichkeiten kann der Dozent seine eigenen Personaldaten im System aktualisieren und eventuelle weitere berechtigte Administratoren (wie z.B. studentische Hilfskräfte oder Sekretärinnen) anmelden.

4 Technische Umsetzung

Die technische Realisation des Portals wird durch die Architektur des Systems beleuchtet. Hierzu wird auf die Datensicht, die Softwarearchitektur, die Softwareausstattung und die Hardwareplattform eingegangen.

4.1 Datensicht

Die Datensicht wird anhand des konzeptionellen und logischen Schemas dargestellt. Das konzeptionelle Schema beschreibt die Basis-Entity-Typen der Datenbank:

- users (Benutzer des Systems)
- fak (Fakultäten)

- lva (Lehrveranstaltungen)
- ls (Lehrstühle)
- mylva (persönliche Lehrveranstaltungen)
- aktuell (aktuelle Nachrichten)
- theme (Thema / Design)
- temp (Temporäre Benutzer)
- bib (Bibliotheken)

Die meisten dieser Entity-Typen sind selbsterklärend. Der Entity-Typ „mylva“ repräsentiert den Zusammenhang zwischen dem jeweiligen Benutzer und den abonnierten Lehrveranstaltungen. Der Typ „users“ enthält sämtliche Benutzerdaten der Dozenten und Studenten. Der Entity-Typ „temp“ dient nur zur temporären Zwischenspeicherung der Anmeldungen von Dozenten. Seine Tabellenstruktur entspricht dem Entity-Typ „users“. Das einstellbare Aussehen des Webinterfaces wird im Entity-Typ „theme“ gespeichert. Das logische Schema der Datenbank ist in Abbildung 7 in Form eines Datenmodells zu sehen:

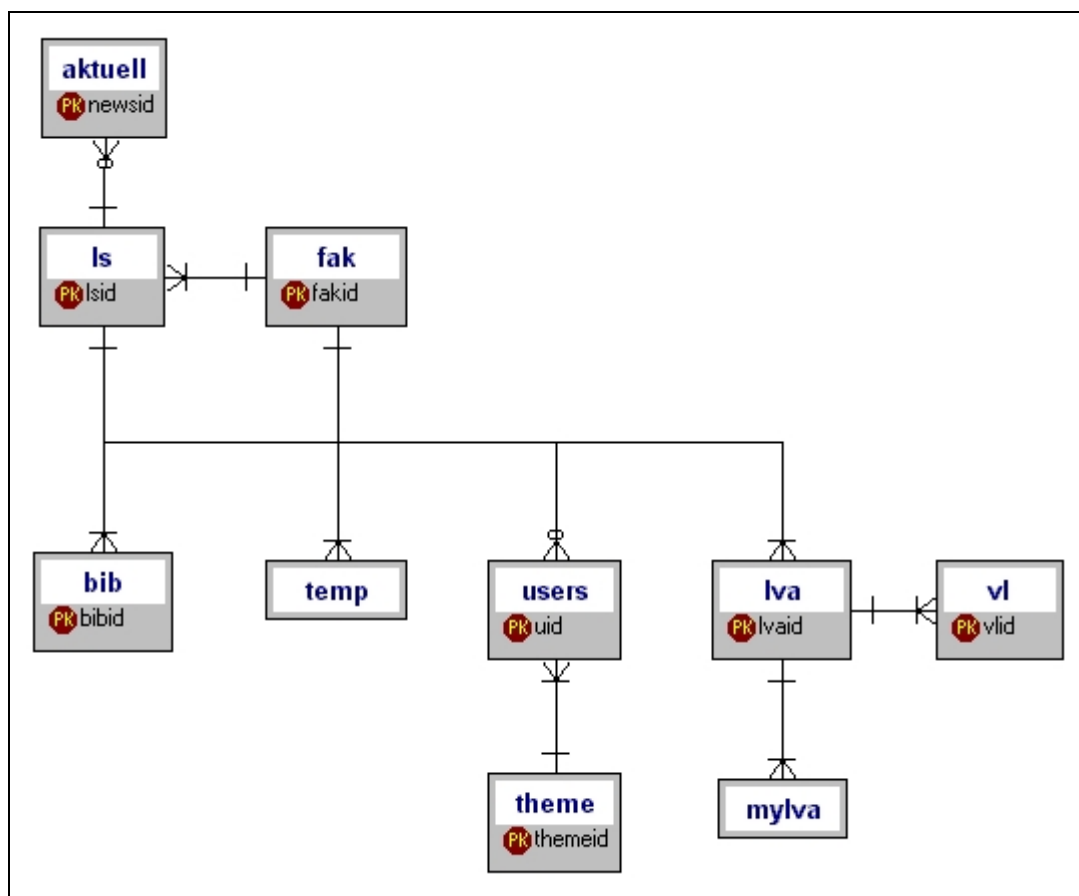


Abbildung 7: Datenmodell der VUR Version 2

4.2 Softwarearchitektur

Die Architektur beschreibt die Struktur des Softwaresystems durch Systemkomponenten und ihre Beziehungen untereinander. Systemkomponenten sind abgegrenzte Teile eines Softwaresystems, die wiederum die Bausteine für die physikalische Struktur einer Anwendung darstellen. [BALZ96, S. 639] Die Lernplattform VUR besteht aus drei Ebenen:

- GUI/Client
- Applikationsserver (mit integrierter Datenbank)
- Medienserver / Kommunikationsserver

In nachfolgender Grafik (Abbildung 8) ist die Architektur des Portals in Form der Komponenten und Beziehungen zwischen diesen dargestellt. [SCHA01, S. 102 u. SCHA02, S. 34]

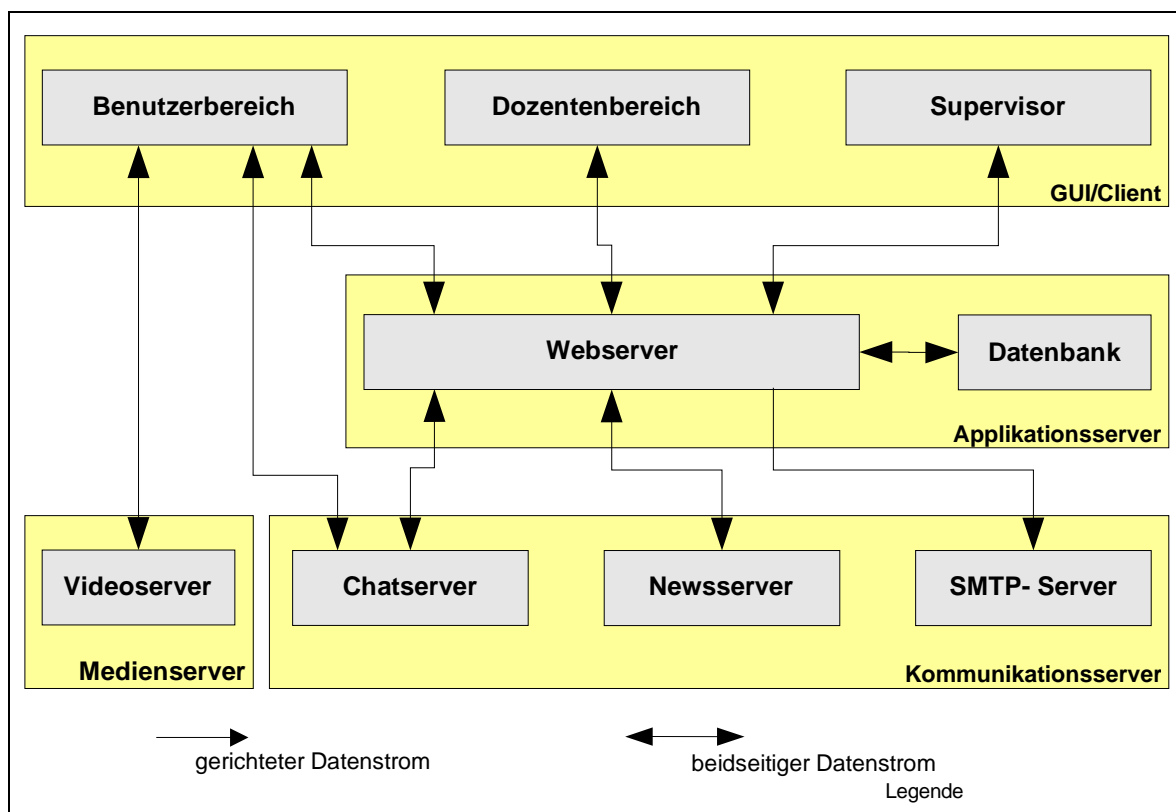


Abbildung 8: Architektur der VUR Version 2

Die erste Ebene stellt die auf den Client übertragenen Webseite dar, die im Browser angezeigt wird. Der Benutzer kann über sie mit dem System in Interaktion treten. Diese Ebene ist nach Benutzergruppen aufgeteilt: Benutzer-, Dozenten- und Supervisorbereich. Letzterer steht jedoch ausschließlich dem Webadministrator selbst zur Verfügung. Der Benutzerbereich kommuniziert über den Webserver in der zweiten Ebene mit den unterschiedlichen Systemkomponenten in der dritten Ebene, den Kommunikationsservern und dem

Medienserver. Der Webserver interagiert mit der integrierten Datenbank und ruft dort die jeweils aktuellen Daten ab bzw. nimmt die vom Dozenten gewünschten Eingaben vor. [SCHA01, S. 103]

4.3 Softwareausstattung

Das Grundsystem der VUR 2 basiert auf einem Webapplikationsserver, der die Daten aus einer relationalen Datenbank abfragt und diese entsprechend aufbereitet im Internet darstellt. Daneben werden für die jeweiligen Kommunikationsdienste noch weitere Serverdienste benötigt. Dazu zählen Newsserver, SMTP-Server, Medienserver und das Chatsystem.

In Tabelle 1 sind die eingesetzten Systeme aufgelistet. Sie wurden für den Einsatz in der VUR Version 1 anhand einer Evaluation [SCHA01, S. 90-99] ausgewählt und sind auch in der VUR 2 in entsprechend neuen Versionen im Einsatz:

Web-/Applikationsserver	ROXEN Webserver 2.2.252 (http://www.roxen.com/)
Datenbanksystem	MySQL 3.22 (integriert) (http://www.mysql.com/)
Newsserver	DNews (mit DNewsweb) V. 5.5g4 (http://www.netwinsite.com/)
Chatserver	SPiN Chatsystem Gold 2.4 (http://www.spinchat.de/)
Mediaserver	REAL System 8 (http://www.realnetworks.com/)
Betriebssystem	Mandrake Linux 8.1 (http://www.linux-mandrake.com/)

Tabelle 1: Softwareausstattung

4.4 Hardwareplattform

Generell hängt die Hardwareplattform, auf der die VUR 2 lauffähig ist, von der zuvor beschriebenen Softwareausstattung ab. Der ROXEN Webserver ist für die Plattformen FreeBSD, HPUX, sämtliche Solaris Varianten, Linux ab Kernel 2.2, MacOS X (Server und Darwin) sowie für Windows NT (4/SP6 bzw. 2000) erhältlich. Der Newsserver DNews der Firma Netwin kann auf sämtlichen Unix-Derivaten einschließlich Linux sowie ebenfalls auf Windows NT/2000 eingesetzt werden. Tabelle 2 zeigt einen Überblick:

	HPUX	AIX	Solaris	Linux	BSD	MacOS X	Win NT
ROXEN	X		X	X	X	X	X
MySQL	X	X	X	X	X	X	X
DNEWS	X	X	X	X	x	X	X
CHAT	X	X	X	X	X	X	X
REALSystem	X	X	X	X	X		X

Tabelle 2: Übersicht über die einsetzbaren Betriebssysteme

Die in Tabelle 2 aufgeführten Betriebssysteme implizieren wiederum die möglichen Hardwareplattformen. HP-UX und AIX sind nur auf den entsprechenden Mainframes der Hersteller Hewlett Packard sowie IBM verfügbar. Solaris ist eine Eigenentwicklung der Firma SUN Microsystems und ist für die SPARC-Plattform (<http://www.sparc.org/standards.html>) sowie eingeschränkt auch für die x86 (Intel) Plattform erhältlich. Linux und BSD (<http://www.netbsd.org/> od. <http://www.freebsd.org/>) laufen auf nahezu jeder verfügbaren Hardwareplattform, d.h. es werden von x86 über SPARC bis zu PPC (Entwicklung von Motorola, Apple und IBM, siehe z.B. <http://www-3.ibm.com/chips/products/powerpc/>), teilweise sogar MIPS (<http://www.mips.com/products/index.html#Anchor-Architecture>) Versionen im Internet zum Download bereitgestellt. Die Skalierung der Hardware an sich hängt sehr stark von der Intensität der Nutzung ab.

5 Resümee und Ausblick

E-Learning-Plattformen und Portale sind mittlerweile auch im deutschsprachigen Raum keine Seltenheit mehr. Die Zielsetzung der VUR 2 liegt weniger in der Möglichkeit einen vollständigen Studiengang anzubieten, sondern in der Ergänzung und Vertiefung des bestehenden Unterrichts. Dazu ist es notwendig, dass nicht nur ein Experten-Team die Inhalte für das Portal aufbereiten kann, sondern jeder Dozent die Möglichkeit hat, Content zu erstellen und diesen auch selbst im Portal verfügbar machen kann. Um jedoch bei den Dozenten auf Akzeptanz zu stoßen, muss das Portal ein einfach zu bedienendes Interface anbieten, das den Dozenten ohne tieferes Computerfachwissen unterstützt. Genau diese Leistung bietet die VUR 2. In einer Zeit, in der Kosten eines der wichtigsten Kriterien für jegliche Entscheidung sind, kann die VUR 2 als Low-Budget-Lösung glänzen. Das gesamte System basiert auf Open-Source-Software und steht selbst unter der GPL (General Public Licence, <http://www.gnu.org/copyleft/gpl.html>). Lediglich der Newsserver und das Chatsystem stehen unter einer sog. Educational-Licence kostenfrei für Bildungseinrichtungen zur Verfügung. Somit entstehen nur Kosten für die Hardwareausstattung, Einrichtung und Wartung des Systems. Aber auch auf Seiten der Studenten bietet die VUR 2 viele Möglichkeiten, die das Studium erleichtern und verbessern. Der Student kann ohne Konfigurationsaufwand von jedem Ort aus über das Portal auf alle elektronischen Lehrmittel zugreifen und die synchronen sowie asynchronen Kommunikationsmittel unter einer einheitlichen Oberfläche nutzen. Durch die Personalisierung kann er jederzeit feststellen, welche Vorlesungen bzw. Lehrmittel seit seinem letzten Login oder einem entsprechenden Zeitraum neu hinzugekommen sind. Dies erleichtert es ihm, den Überblick über seine Materialien und den Lernfortschritt zu

halten. In Zukunft wird auch das ortsunabhängige Mobile-Learning mit Hilfe von PDAs und Handys immer mehr an Bedeutung gewinnen. Dieser Entwicklung wird schon jetzt Rechnung getragen, indem die Mobile-Dienste schrittweise in das Portal integriert werden. Verfügbar sind derzeit schon u.a. ein Abruf von aktuellen Mitteilungen, der Download von Lehrmaterialien (PDF und MP3) für PDAs und der Abruf von Kontakt-Informationen der Dozenten.

Die E-Learning-Plattform VUR 2 wird ständig weiterentwickelt, um auf die Wünsche und Anregungen der Anwender einzugehen. Die Resonanz von Seiten der Fakultäten und Lehrstühle ist sehr gut. Es sind aktuell (Stand Juli 2002) zehn Fakultäten (von zwölf) und insgesamt 31 Lehrstühle aktiv. Von Seiten der Studenten, die das Portal aktiv nutzen (derzeit sind 817 Benutzer personalisiert, wobei eine Nutzung auch ohne Personalisierung möglich ist), sind ebenfalls durchweg positive Rückmeldungen vorhanden. Um dieses Feedback durch Zahlen entsprechend belegen zu können und um die Weiterentwicklung in die richtige Richtung zu lenken, wird ab dem Wintersemester 2002 im Rahmen einer Diplomarbeit an einem Pädagogik Lehrstuhl der Universität eine Evaluation unter den Studenten durchgeführt.

6 Literaturverzeichnis

- [BALZ96] Balzert, H.: Lehrbuch der Software-Technik: Software Entwicklung. Heidelberg, Berlin, Oxford 1996.
- [LEHN00] Lehner, F.: Chancen und Grenzen der Virtualisierung an Hochschulen – Bestandsaufnahme und Bewertung der Situation in Deutschland. Bericht Nr. 37, Universität Regensburg, Schriftenreihe des Lehrstuhls für Wirtschaftsinformatik III, 2. Auflage, 2000.
- [LEHN01] Lehner, F.: Teleteaching in der Wirtschaftsinformatik - Erfahrungen mit einem Teleseminar an der Universität Regensburg. Bericht Nr. 25, Universität Regensburg, Schriftenreihe des Lehrstuhls für Wirtschaftsinformatik III, 6. Auflage, 2001.
- [LEHN02] Lehner, F.: E-Learning - Virtueller Unterricht über das Internet am Beispiel von Hochschulen und Universitäten. Bericht Nr. 52, Universität Regensburg, Schriftenreihe des Lehrstuhls für Wirtschaftsinformatik III, 3. Auflage, 2002.
- [SCHA00] Schäfer, K. J.: Eine Plattform für Learning on Demand im Internet am Beispiel der Virtuellen Universität Regensburg. Bericht Nr. 42, Universität Regensburg, Schriftenreihe des Lehrstuhls für Wirtschaftsinformatik III, 2000.
- [SCHA01] Schäfer, K. J.: Internetportal für eine Virtuelle Hochschule Entwicklung einer Plattform für Learning on Demand. Wiesbaden 2001.
- [SCHA02] Schäfer, K.J.: Virtuelle Universität Regensburg - Version 2. Bericht Nr. 58, Universität Regensburg, Schriftenreihe des Lehrstuhls für Wirtschaftsinformatik III, 2002.



Netzgestützte Lehre - Vom Projekt zum Regellehrbetrieb

Folker Caroli, Christa Hauenschild, Daniela Wagner

Institut für Angewandte Sprachwissenschaft
Universität Hildesheim
Marienburger Platz 22
31141 Hildesheim
caroli@rz.uni-hildesheim.de

1 Einleitung

Der Einsatz der Informations- und Kommunikationstechnologie für die Entwicklung neuer Lehrformen der mediengestützten Lehre innerhalb der Hochschulen wird aktuell mit erheblichen Mitteln gefördert. In zahlreichen Projekten sind innovative Wege der Nutzung dieser Technologien für die Verbesserung der Lehre an den Hochschulen entwickelt worden. Allerdings folgt aus dem erfolgreichen Abschluss eines Projekts nicht zwangsläufig, dass die Ergebnisse, die in ihm erzielt wurden, auch nachhaltig im Regellehrbetrieb der Hochschulen genutzt werden. Dieses Problem wurde in zahlreichen Beiträgen auf der Jahrestagung 2001 der Gesellschaft für Medien in der Wissenschaft¹ diskutiert. Michael Kerres (Kerres 2001) weist zu Recht darauf hin, dass für die Verankerung medien- und netzgestützter Formen der Lehre Veränderungen in der Hochschulorganisation nötig sind, die sehr stark das Selbstverständnis universitärer Lehre berühren.

Das Ende 2001 abgeschlossene Projekt „Virtueller Campus Sprache – Wissen – Information“ hat von Anfang an die Untersuchung der Bedingungen für die Übernahme der Projektergebnisse in den Regellehrbetrieb als ein Projektziel verfolgt. Die Idee war dabei, diese Frage nicht auf einer zu allgemeinen Ebene zu diskutieren, sondern zu versuchen, in dem Prozess der Entwicklung sinnvoller Formen der netzgestützten Lehre zugleich Klarheit darüber zu gewinnen, welche Bedingungen erfüllt sein müssen, damit diese Formen der Lehre im Regellehrbetrieb genutzt werden können.

¹ Vgl. die Beiträge der Sektion Struktur und Implementierung im Tagungsband. (Wagner/Kindt 2001).



In dem folgenden Beitrag werden solche Überlegungen für die in dem Projektteil des IfAS (Institut für Angewandte Sprachwissenschaft) entwickelte Lehrform des virtuellen Seminars vorgestellt. Diese Überlegungen sind abgeleitet aus umfangreichen Arbeiten zur Evaluation dieser Lehrform, die im Zusammenhang mit dem parallel im IfAS bearbeiteten Projekt „Lernen im Netz“ vorgenommen wurden².

2 Das Konzept des virtuellen Seminars

2.1 Projektkontext

Das Projekt „Virtueller Campus“ war ein Verbundprojekt der folgenden Partner:

- Institut für Angewandte Sprachwissenschaft (IfAS), Universität Hildesheim mit den Studiengängen Internationales Informationsmanagement und Internationale Fachkommunikation
- Institut für Rechnergestützte Wissensverarbeitung (KBS), Universität Hannover mit dem Studiengang Elektrotechnik/ Informatik
- Institut für Semantische Informationsverarbeitung (ISIV), Universität Osnabrück mit den Studiengängen Computerlinguistik/Künstliche Intelligenz und Cognitive Science
- Zentrum für Fernstudium und Weiterbildung (ZFW), Universität Hildesheim

Die Institute, die in diesem Projektverbund zusammenarbeiten, betreuen unterschiedliche Studiengänge, die aber inhaltliche Überschneidungen aufweisen, die den Austausch von Lehrveranstaltungen zwischen den Projektpartnern interessant erscheinen lassen. Jeder der Projektpartner kann Lehrveranstaltungen anbieten, die als Ergänzung und Vertiefung des Lehrangebots der anderen Partner interessant sind, aber dort nicht oder nicht in der gleichen inhaltlichen Breite angeboten werden können. Diese Form des Projektverbunds steckt einen Kooperationsrahmen ab, welcher der Zusammenarbeit in der Lehre über die Projektarbeit hinaus ein organisatorisches Fundament bieten kann.

Über einen Zeitraum von viereinhalb Jahren verfolgte der Projektverbund die folgenden Ziele:

² Die ausführliche Darstellung der Methode und der Ergebnisse dieser Evaluation ist Gegenstand der Habilitationsschrift von Christiane Schmidt (Schmidt in Vorbereitung).

- Entwicklung und Verknüpfung intelligenter Lehr- und Lern-Systeme für einen stabilen Einsatz in der Hochschullehre, die auch für eine Anwendung außerhalb des Projektverbunds geeignet sein sollten
- Entwicklung fachlicher Lehrmodule mit den Schwerpunkten Computerlinguistik, Künstliche Intelligenz, objektorientierte und logische Programmierung, Cognitive Science, Multilinguale Informationssysteme
- Klärung der Voraussetzungen des Einsatzes der Module im Regellehrbetrieb
- Beiträge zur Hochschulentwicklung für die Lösung der Probleme im Zusammenhang mit dem Einsatz von virtuellen Lehr- und Lernformen im Regellehrbetrieb

Das virtuelle Seminar war ein wesentlicher Bestandteil des am IfAS durchgeführten Teilprojekts. Im Zentrum dieses Teilprojekts stand die Entwicklung und Erprobung des Konzepts der verteilten Kooperation in virtuellen Seminaren.

2.2 Didaktisches Konzept des virtuellen Seminars

Das didaktische Konzept des virtuellen Seminars ist an anderen Stellen ausführlich beschrieben worden (Caroli 2000, Caroli /Zandner 2001, Caroli / Hauenschild / Hellmann 2001). Wir beschränken uns deshalb hier darauf, die Kernpunkte zusammenzufassen.

Studieren in verteilter Kooperation

Mit dem Konzept der verteilten Kooperation lässt sich eine grundlegende Veränderung in weiten Bereichen der Arbeitswelt beschreiben (vgl. Caroli et al. 1995, van Haaren 1997, Picot 1997, Reichwald et al. 1998). Die Betriebsstätte, in der die ArbeitnehmerInnen zusammenkommen und in der alle nach festgelegten Zuständigkeiten die anfallenden betrieblichen Aufgaben abarbeiten, verliert als strukturierendes Moment der Arbeitsorganisation an Bedeutung. Arbeitsabläufe werden zunehmend projektorientiert organisiert. Dies bedeutet, dass für eine begrenzte Aufgabe ein Team zusammengestellt wird, das nur für die Erledigung dieser Aufgabe Bestand hat. Es ist dabei nicht mehr entscheidend, dass die Mitglieder des Projektteams ihre Aufgaben alle an einem Ort erledigen. Die Möglichkeiten der netzgestützten Kommunikation und Kooperation werden zur standortübergreifenden und zeitlich flexibilisierten Zusammenarbeit genutzt.

Gerade in den Berufsfeldern, für die die Studiengänge der beteiligten Projektpartner qualifizieren, gewinnt diese Form der Arbeitsorganisation zunehmend an Bedeutung. Es wird deshalb für die Ausbildung der Studierenden dieser

Bereiche wichtig, solche Arbeitsformen einzuüben und die Prozesse und Probleme dieser Form der Projektarbeit zu reflektieren. Die Schlüsselqualifikationen, die dabei zu vermitteln sind, liegen nicht allein in der Beherrschung der für die verteilte Kooperation eingesetzten Technik. Wichtiger noch ist die Entwicklung einer spezifischen kommunikativen und sozialen Kompetenz. Gerade die für Projektarbeit als zentral angesehenen Schlüsselqualifikationen wie Teamfähigkeit, Kooperationsfähigkeit, Konfliktfähigkeit und Kommunikationsfähigkeit verlangen in der verteilten Kooperation eine jeweils spezifische Ausprägung.

Unter dieser Perspektive haben wir unser Konzept eines virtuellen Seminars entwickelt. Es soll den Studierenden die Möglichkeit eröffnen, Formen der verteilten Kooperation selbst zu erproben und zu reflektieren. Dabei greifen wir vielfältige Überlegungen zum kooperativen und explorativen Lehren und Lernen auf (siehe hierzu die Beiträge der Tagung "Computergestütztes Kooperatives Lernen (D-CSCL 2000)" (Uellner/Wulf 2000, Schulmeister 1997, S. 76-81, Klimsa 1993, Mandl et al. 1997, Gräsel 1997).

Inhalte

Die Bedingungen der verteilten netzgestützten Kooperation lassen sich sehr schwer zu Ausbildungszwecken simulieren; sie brauchen eine sachlich und inhaltlich ausgewiesene Motivation. In dem virtuellen Seminar wird deshalb Studierenden des Hauptstudiums aus den unterschiedlichen Studiengängen der beteiligten Projektpartner die Möglichkeit geboten, einen Gegenstandsbe- reich aus den Themengebieten der maschinellen Sprachverarbeitung, der ma- schinellen und maschinengestützten Übersetzung und der multilingualen In- formationssysteme gemeinsam zu erarbeiten.

In Anlehnung an die Arbeitsweise realer Forschungs- und Entwicklungspro- jekte sollen dabei die unterschiedlichen Sichtweisen und die unterschiedlichen Kompetenzen und Kenntnisse der Studierenden für die Seminararbeit frucht- bar gemacht werden. Die Aufgabenstellungen haben in der Regel einen An- wendungsbezug und sind dabei hinreichend komplex gewählt, so dass sich die Notwendigkeit ergibt, für ihre Lösung die unterschiedlichen Kenntnisse und Fähigkeiten der Studierenden zu nutzen.

Daraus folgt, dass die einzelnen Aufgaben nicht ohne weiteres von lokalen Gruppen der beteiligten Projektpartner bearbeitet werden können, sondern dass es von Vorteil ist, hierfür standortübergreifende Arbeitsgruppen zu bil- den. Somit ergibt sich aus der Aufgabenstellung eine inhaltliche Begründung

für die projektorientierte Arbeitsform und für die Notwendigkeit der verteilten Kooperation innerhalb des Seminars.

Wenn für eine Lehrveranstaltung die organisatorischen und technischen Voraussetzungen für eine netzgestützte verteilte Kooperation geschaffen werden, eröffnet dies eine weitere Dimension der inhaltlichen Arbeit. Diese Infrastruktur kann auch dafür genutzt werden, externe ExpertInnen, die auf Anwendungsfeldern der in dem Seminar zu erarbeitenden Fachinhalte tätig sind, direkt in die Seminararbeit einzubinden. Hierin liegt ein weiterer Gewinn für die Universitätslehre und eine zusätzliche Motivation für die Studierenden.

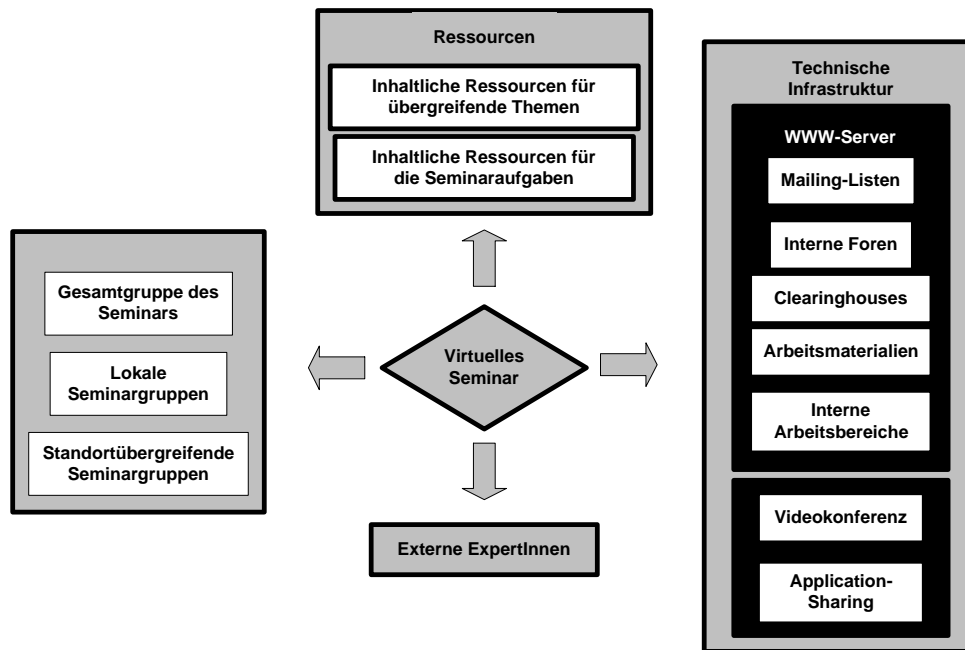
Arbeitsformen

Unter den Bedingungen der verteilten Kooperation verändern sich die Rollen von Lehrenden und Studierenden innerhalb des Seminars. Wenn Projektarbeit in verteilter Kooperation tatsächlich erprobt werden soll, bedeutet dies, dass den studentischen Arbeitsgruppen größere Verantwortung für die Planung und Organisation ihrer Arbeit und damit für das Erreichen eines eigenen Arbeitsergebnisses übertragen werden muss. Die Lehrenden nehmen in diesem Arbeitsprozess in hohem Maße eine orientierende, beratende und moderierende Funktion wahr.

Aufbauend auf den Erfahrungen mit der Arbeitsform von Forschungs- und Entwicklungsprojekten in verteilter Kooperation sieht unser Konzept einen planvollen Wechsel zwischen Präsenzphasen und netzgestützten synchronen und asynchronen Arbeitsformen vor, so wie sie in der Realität verteilter Kooperation vorzufinden sind.

Lernumgebung

Aus diesen didaktischen Überlegungen lassen sich wesentliche Elemente der Lernumgebung des Seminars ableiten. Im Gegensatz zu vielen Ansätzen zur Organisation von virtuellen Lehr- und Lernformen verstehen wir dabei unter Lernumgebung nicht nur das technische Arrangement, mit dem die für eine Lehrveranstaltung als sinnvoll erachteten Formen der Kommunikation und Informationsvermittlung realisiert werden (vgl. hierzu: Lehren und Lernen mit neuen Medien 1999). Eine Lernumgebung umfasst für uns vielmehr die Gesamtheit der sozialen, organisatorischen, inhaltlichen und technisch/materiellen Faktoren, die in planvoller Weise kombiniert werden, um die intendierte Form des Lernens und Lehrens zu unterstützen. Die Grafik „Lernumgebung für virtuelle Seminare“ fasst zusammen, welche Elemente in der Lernumgebung für ein virtuelles Seminar integriert sind.



Lernumgebung für virtuelle Seminare

Als wichtiges Hilfsmittel für die Unterstützung der Arbeit der Seminargruppen werden unterschiedliche inhaltliche Ressourcen bereitgestellt. Entsprechend dem Konzept des explorativen Lernens, das dem Seminar zu Grunde liegt, bestehen diese Ressourcen nicht ausschließlich in didaktisch aufbereiteten Lehr-Lernmodulen zu dem Seminargegenstand, sondern enthalten in großem Umfang auch Materialien, die von den Arbeitsgruppen selbst auf ihre Relevanz für die Lösung der gestellten Aufgabe hin ausgewertet werden sollen. Diese Ressourcen werden in der Form eines Clearinghouses zu den für die Seminararbeit relevanten Themengebieten aufbereitet. Dieses Clearinghouse wird mit jedem Seminardurchgang erweitert und aktualisiert, so dass eine ständig wiederverwendbare inhaltliche Ressource entsteht.

Zusätzlich zu den inhaltlichen Ressourcen werden Anleitungen und Hinweise zur Gruppenarbeit in netzgestützter verteilter Kooperation und zur effektiven und zielgerichteten Nutzung der netzgestützten Kommunikationsdienste angeboten. Erst unter dieser erweiterten Perspektive der Lernumgebung lässt sich eine begründete Wahl der eingesetzten technischen Infrastruktur vornehmen. In den virtuellen Seminaren kommen sowohl ein integrierter Server als auch ein Videokonferenzsystem zum Einsatz. Als integrierten Server haben wir die auf dem Hyperwave-Server aufbauende eLearning Suite (vormals Gentle-WBT) benutzt. Hier sind alle Kommunikations- und Informationsdienste integriert, die für das Seminar benötigt werden. Zum einen können in der eLearning Suite alle für das Seminar relevanten inhaltlichen Ressourcen bereitgestellt werden. Zum anderen erlaubt diese Plattform auch eine flexible

Einrichtung von Mailinglisten, Diskussionsforen und Chat-Räumen entsprechend der Gruppenstruktur des Seminars. Weiterhin ist die Vergabe von abgestuften Zugriffsrechten möglich, so dass für jede Arbeitsgruppe ein eigener virtueller Raum eingerichtet werden kann, der nur den Mitgliedern dieser Arbeitsgruppe zugänglich ist. Dies ist eine wichtige Voraussetzung für den Austausch und die Diskussion von Beiträgen im Arbeitsprozess der Gruppen.

Ein wichtiges Element der Lernumgebung ist die Möglichkeit externe ExpertInnen, die auf Anwendungsfeldern, der in dem Seminar zu erarbeitenden Fachinhalte tätig sind, direkt in die Seminararbeit einzubinden. Hierin liegt ein weiterer Gewinn für die Universitätslehre und eine zusätzliche Motivation für die Studierenden. Diese ExpertInnen werden mit ihren Arbeitsschwerpunkten auf den Seminarseiten vorgestellt, so dass eine direkte Kontaktaufnahme möglich wird.

Als Videokonferenzsystem haben wir sowohl Netmeeting als auch Mbone erprobt. Beide Systeme erlauben mit unterschiedlichen Einschränkungen und Schwächen die Durchführung von Videokonferenzen zwischen mehreren Standorten und mit unterschiedlichen Plattformen. Beide stellen Tools für die Durchführung von Videokonferenzen, für das Application Sharing sowie für zusätzliche Kommunikationsmittel wie ein Whiteboard und Chat zur Verfügung.

Die hier beschriebene organisatorische und technische Infrastruktur schafft die nötigen Voraussetzungen für eine netzgestützte verteilte Kooperation innerhalb des Seminars.

3 Erprobung und Evaluation des Konzepts

Das Konzept des virtuellen Seminars wurde während der gesamten Laufzeit des Projekts in drei Durchgängen erprobt und evaluiert. Die Evaluation der ersten beiden Durchgänge wurde in Zusammenarbeit mit dem Projekt „Lernen im Netz“ durchgeführt. Die Ergebnisse der Evaluation der einzelnen Durchgänge konnten zur Optimierung des Konzepts für die nachfolgenden Durchgänge genutzt werden. (Vgl. Schmidt 2000, Schmidt in Vorbereitung). Der dritte Seminardurchgang war insofern den Bedingungen des Einsatzes im Regellehrbetrieb angenähert, als mit dem Institut für Sprachwissenschaften der Universität Regensburg ein Partner außerhalb des Projektverbundes beteiligt wurde. Aufgrund der stark divergierenden Zeitpunkte des Semesterbeginns der an dem Seminar beteiligten Partnerhochschulen war es in diesem Durchgang nicht möglich, einen Startworkshop durchzuführen. Die projekt-

orientierte Arbeitsweise wurde deshalb in diesem Durchgang etwas zurückgenommen. Die Beiträge der Studierenden, die zu gleichen Teilen durch lokale und durch standortübergreifende Arbeitsgruppen erarbeitet wurden, waren stärker der Form von Referaten in einem Hauptseminar angenähert. Ein weiterer wichtiger Gesichtspunkt war, dass das Zeitbudget, das die für die inhaltliche Durchführung des Seminars verantwortlichen Dozentin in die Seminararbeit eingebracht hat, den Zeitaufwand für die Durchführung eines Hauptseminars nicht wesentlich überschreiten sollte.

Auch für diesen Durchgang wurde eine Evaluation auf der Basis einer Fragebogenerhebung und von Gruppendiskussionen am Ende des Seminars durchgeführt. Die Ergebnisse der Evaluation sind in dem Abschlussbericht des Gesamtprojekts zugänglich (Virtueller Campus 2002).

Die Ergebnisse der Evaluation zeigen, dass das Konzept des virtuellen Seminars insgesamt tragfähig ist und auf eine starke Akzeptanz bei den Studierenden stößt. Das virtuelle Seminar wird als interessante und sinnvolle Erweiterung des Lehrangebots angesehen. Die Möglichkeit, sich mit Formen der netzgestützten verteilten Kooperation vertraut zu machen und Schlüsselqualifikationen für diese Arbeitsform zu erwerben, wird positiv bewertet. Die Attraktivität dieser Lehrform wird noch höher bewertet, wenn das virtuelle Seminar Inhalte zum Gegenstand hat, die an der eigenen Universität nicht oder nur eingeschränkt vermittelt werden können.

Allerdings wird in den Aussagen der Studierenden deutlich, dass virtuelle Lehr- und Lernformen als ausschließliche Form des Studierens nicht als angemessen angesehen werden. Dies bedeutet, dass die Studierenden der Vision einer vollständig virtuellen Hochschule skeptisch gegenüberstehen.

Bei einer insgesamt der positiven Bewertung des virtuellen Seminars zeigen die Ergebnisse der Evaluation eine Reihe von Gesichtspunkten auf, die für die Übernahme des Konzepts in den Regellehrbetrieb berücksichtigt werden sollten.

3.1 Technikebene

Während des Projekts wurden für die Vermittlung der synchronen und asynchronen netzgestützten Kommunikation unterschiedliche technische Konfigurationen erprobt. Hierbei waren Aspekte der Verfügbarkeit, des Handlings und der Robustheit der Technik wichtig. Bei der Übernahme des Konzepts des virtuellen Seminars in den Regellehrbetrieb ist sicherzustellen, dass die

benötigte Infrastruktur dauerhaft aufgebaut und fortentwickelt wird. Hierfür sind die notwendigen Investitionsmittel und personellen Ressourcen bereitzustellen. Aufbau und Wartung der Infrastruktur kann nicht den Lehrenden als zusätzliche Aufgabe übertragen werden; andernfalls wird die mit der virtuellen Lehre verbundene zusätzliche Belastung unverhältnismäßig hoch.

Eine zufriedenstellende Qualität der Audio-Video-Übertragung ist eine unabdingbare Voraussetzung für die Gestaltung der netzgestützten synchronen Arbeitsphasen des virtuellen Seminars; andernfalls kommt es zu einer erheblichen Belastung für die Arbeit des Seminarplenums. Durch den Einsatz spezieller Hardware für die Vermittlung von Videokonferenzen im Forschungsnetz kann ein solcher Standard mit vertretbaren Kosten erreicht werden. Dabei ist eine Standardisierung der Ausstattung der in einem Lehrverbund beteiligten Partner anzustreben.

Weiterhin sollte verstärkt auf eine adäquate Ausstattung der für die Videokonferenz genutzten Seminarräume hingewirkt werden, die eine effektive Regie und Moderation der Videokonferenz unterstützt (Sitzanordnung, mehrere Kameras mit Regiepult, Mikrofonanlage, Projektion, Raumbeschallung, Ausleuchtung des Raumes).

Während der Projektlaufzeit war bei den Studierenden der Studiengänge, für die das Angebot des virtuellen Seminars geschaffen wurde, eine deutliche Entwicklung der Kenntnisse im Umgang mit den Kommunikationsdiensten des Internets sowie der privaten Verfügbarkeit von Computern und des privaten Zugangs zum Internet festzustellen. Zu Beginn des Projekts stellten mangelnde Kenntnisse im Umgang mit dem Computer und insbesondere mit den Internetdiensten sowie das Fehlen eines privaten Internetzugangs noch ein Problem für die Arbeit im virtuellen Seminar dar. Im letzten Durchgang schätzten die TeilnehmerInnen ihre Vorkenntnisse im Umgang mit Computern und in der Nutzung der Internetdienste überwiegend als gut bis sehr gut ein. Weiterhin verfügten alle TeilnehmerInnen über einen privaten Computer mit Internetzugang. Damit waren sehr gute Voraussetzungen für die Nutzung der Web-basierten Lernplattform (eLearning Suite von Hyperwave) gegeben. Dies wird dadurch bestätigt, dass in diesem Seminardurchgang von den TeilnehmerInnen der Aufwand für die Einarbeitung in die eLearning Suite überwiegend als angemessen betrachtet wurde und keine schwerwiegenden technischen Schwierigkeiten bei der Benutzung aufgetreten sind.

Für die Lösung der technischen Probleme sollte an jedem Standort des virtuellen Seminars für entsprechenden technischen Support gesorgt werden. Als

wichtiges Ergebnis der Evaluation kann festgehalten werden, dass es sinnvoll ist, die Studierenden selbst an der Lösung der Probleme zu beteiligen. Auf diese Weise können die verschiedenen technischen Fertigkeiten, welche die Studierenden in die Seminararbeit einbringen, für die Arbeit des gesamten Seminars genutzt werden. Dies wird auch von den Studierenden gewünscht, wie die folgende Aussage aus einer Diskussion während des Abschlussworkshops, die beispielhaft für andere stehen kann, belegt:

- Es war schade, dass die Studierenden nicht bei technischen Fragen einbezogen wurden. Z.B. könnte ein Workshop zu Beginn dazu dienen, dass sich die Teilnehmer des Seminars mit der Technik vertraut machen und ihr Wissen zusammenführen. Die Pausen, die bei den technischen Problemen entstanden, waren eine nutzlose Zeit für die Teilnehmer.

3.2 Didaktikebene

Die Evaluation der drei Seminardurchgänge zeigt, dass das didaktische Konzept sowohl für die Vermittlung von Schlüsselqualifikationen als auch für die Erarbeitung der fachlichen Inhalte tragfähig ist. Die überwiegende Zahl der TeilnehmerInnen betrachtet die Teilnahme an einem solchen Seminar als eine wertvolle Erfahrung und würde an einem solchen Seminar wieder teilnehmen, wenn die notwendigen technischen und organisatorischen Voraussetzungen für die Arbeit in verteilter Kooperation gegeben sind.

Selbstorganisation der Studierenden

Das Konzept des Studierens in verteilter Kooperation gibt den Arbeitsgruppen des Seminars eine größere Autonomie in der Organisation ihrer Arbeit. Diese Selbstorganisation der Arbeitsgruppen stellt an die Studierenden hohe Anforderungen. Dabei ist weniger die Aufteilung der inhaltlichen Aufgaben ein Problem als vielmehr die Koordination und die Integration der Teilergebnisse der Arbeit. Absprachen in der verteilten Kooperation sind mühsamer zu treffen. Die informellen Kontakte, die sich in Situationen der Präsenzlehre üblicherweise ergeben und die Gruppenarbeit stützen und absichern, fallen in der netzgestützten verteilten Kooperation weg. Dies hat zur Folge, dass die Gruppen ihre Arbeit vorausschauend planen und einen genauen Zeitplan erarbeiten müssen. Verabredungen innerhalb eines solchen Arbeitsplanes haben eine strenge Verbindlichkeit und stellen hohe Anforderungen an die Selbstdisziplin jedes einzelnen Mitglieds in der Gruppe.

In diesem Rahmen gewinnt für die Studierenden die Nutzung der Kommunikationsdienste des Internet eine neue Qualität. Die Handhabung der Technik bereitet den Studierenden dabei in der Regel kaum Schwierigkeiten. Die

Probleme der Nutzung sind eher auf der kommunikativen Ebene anzusiedeln. Es bereitet den Studierenden häufig Schwierigkeiten, die einzelnen Formen der netzgestützten Kommunikation in einer der zu lösenden Aufgabe angemessenen Weise zu nutzen.

Diese Schwierigkeiten sind den Studierenden zu Beginn der Seminararbeit nicht bewusst. Eine wichtige Voraussetzung der erfolgreichen Durchführung virtueller Seminare im Regellehrbetrieb besteht deshalb darin, entsprechende Hilfestellung und Anleitung bereitzustellen; nur auf diese Weise kann sichergestellt werden, dass die inhaltliche Arbeit nicht durch technische und organisatorische Probleme in den Hintergrund gedrängt wird.

Die in dem Konzept des virtuellen Seminars vorgesehenen Präsenzphasen in der Form von Workshops zu Beginn und am Ende des Seminars sind für die Selbstorganisation der Arbeitsgruppen und für die Qualitätssicherung der Arbeitsergebnisse der Gruppen von großer Bedeutung. Dies haben in der Evaluation aller drei Durchgänge alle Studierenden hervorgehoben. Dies belegt der folgende Ausschnitt aus einer Gruppendiskussion, die im Rahmen der Evaluation des dritten Durchgangs des virtuellen Seminars bei dem es wie gesagt keinen Startworkshop gab, geführt wurde:

- Ich denk, das ist auch das Problem, dass andere sich ausgeklinkt hatten, das ist, weil man sich nicht sieht und sich nicht kennt, und sich nicht über den Weg läuft, fühlt man sich nicht so verpflichtet. Und das ist anders, wenn man sich 'ne Mail schreibt und sagt, mach mal was, ist es etwas anderes oder ob man sich irgendwann trifft und da sitzt dann einer, der hat immer noch nichts gemacht. Das ist vom Zusammenarbeiten schon was anderes, da macht jeder seinen Teil und wirft es halt irgendwann in einen Topf. (Zustimmung der gesamten Gruppe)
- Daher wäre meiner Meinung nach auch ein Startworkshop wichtig. Es geht gar nicht mal so um den Inhalt, sondern um die persönliche Kontaktaufnahme.
- Man hat dann ein Gesicht im Kopf und verbindet etwas mit dem anderen Menschen.

Die persönliche Kontaktaufnahme zu Beginn des Seminars wird also für die Zusammenarbeit in den Arbeitsgruppen als überaus wichtig und förderlich angesehen. Ebenso war zu beobachten, dass die Gruppen nach der Präsentation und Diskussion ihrer Arbeitsergebnisse in einem Abschlussworkshop ihre Arbeiten einer zum Teil umfangreichen Überarbeitung unterzogen haben.

Rolle der Seminarleitung

Der größeren Autonomie der Studierenden entspricht eine veränderte Rolle der Seminarleitung. Sie legt den inhaltlichen Rahmen der Seminararbeit fest und zeigt Fragestellungen auf, deren Bearbeitung besonders lohnend erscheint; diese orientierende Funktion ist für den Einstieg in die Arbeit von großer Bedeutung. Doch sollen diese inhaltlichen Vorgaben den Arbeitsgruppen Raum lassen, eigenständige Fragestellungen zu entwickeln und eigene Wege der Untersuchung zu gehen. Sie gibt damit einen Teil der inhaltlichen wie der organisatorischen Verantwortung für die Seminararbeit an die Studierenden ab. Zugleich hat sie aber eine moderierende Funktion wahrzunehmen, um den Prozess der Selbstorganisation der Arbeitsgruppen zu unterstützen. Eine wichtige Funktion der Moderation liegt darin, sicherzustellen, dass die Arbeitsgruppen in dem gesetzten Zeitrahmen ein Arbeitsergebnis erreichen können. Hierzu sind klare orientierende Vorgaben hinsichtlich der Aufgabenstellungen und des Zeitplans des Seminars zu geben, sowie geeignete Hilfestellung bei der Bewältigung der organisatorischen Probleme zu leisten.

Als ein Indikator dafür, dass diese Rollenverschiebung nicht immer ideal gelingt, kann die nur zögerliche Nutzung des allgemeinen Diskussionsforums des Seminars in allen drei Semindurchgängen angesehen werden. Die Studierenden fühlen sich anscheinend immer noch von der Seminarleitung unter Druck gesetzt. Ein Beispiel aus der Gruppendiskussion zum dritten Semindurchgang legt diese Sichtweise nahe:

- Aber es war auch so, das man dachte, jetzt muss man etwas reinschreiben. Man schreibt nur etwas rein, damit der Namen präsent ist. Da ist es eher interessant, dass man verschiedene Foren zu den Themen hat.
- In einer face-to-face-Diskussion entwickelt man mehr Emotionen. Im virtuellen Raum sind nach dem letzten Beitrag wieder eine Stunde oder Tage vergangen und man überlegt sich erst, ob man wirklich eine Antwort hat. Wenn man nicht ganz sicher ist, stellt man sie meistens nicht rein.

Diese Sichtweise wird dadurch bestätigt dass, gruppenspezifischen Diskussionsforen von den Arbeitsgruppen zum Teil intensiv genutzt wurden.

Organisationsebene

Das Virtuelle Seminar bedeutet sowohl für Lehrende als auch für Studierende einen erheblichen Mehraufwand. Folgende Aspekte sind wichtig, um diesen Mehraufwand bei der Übernahme des Konzepts in den Regellehrbetrieb zu begrenzen:

- Der Wechsel von Präsenzphasen, synchronen und asynchronen netzgestützten Arbeitsphasen macht eine genaue und vorausschauende Termin-

planung für das gesamte Seminar nötig, so dass sich alle Beteiligten frühzeitig auf diese Termine einstellen können. In der Terminplanung sind zum Teil banale Probleme zu lösen, z. B., dass Beginn und Dauer der Vorlesungszeit an den beteiligten Hochschulen unterschiedlich sein können, auch wenn sie in demselben Bundesland liegen. Die Arbeit in einem virtuellen Seminar benötigt einen von reinen Präsenzveranstaltungen deutlich unterschiedenen Arbeitsrhythmus. Die Studierenden brauchen die vorlesungsfreie Zeit für die Fertigstellung ihrer Arbeiten und benötigen in dieser Zeit weitere Beratung durch die Lehrenden. Dies muss bei der Terminplanung berücksichtigt werden.

- Die Arbeit in verteilter Kooperation und die aufgabenorientierte Nutzung der Kommunikationsmöglichkeiten des Internet sind für die Studierenden in der Regel nicht vertraut. Innerhalb des Seminars sind Materialien und Hilfestellung vorzusehen, die den Studierenden für die Bewältigung dieser Aufgaben eine sichere Orientierung geben.
- Die Abhaltung virtueller Seminare ist eine arbeitsteilige Aufgabe. Zwischen den Partnerhochschulen müssen die Verteilung der Aufgaben (inhaltliche Leitung des Seminars, Betreuung der Seminargruppen an den beteiligten Standorten, technischer Support) in der Planungsphase festgelegt werden.
- Die Bereitstellung der inhaltlichen Ressourcen für die Seminararbeit erfordert in der Regel einen höheren Arbeitsaufwand, als er für Präsenzseminare zu leisten ist. Dieser Aufwand kann aber auf ein vertretbares Maß zurückgeführt werden, wenn die für einen Gegenstandsbereich aufgebauten inhaltlichen Ressourcen so angelegt werden, dass sie für mehrere Semindurchgänge mit unterschiedlicher thematischer Ausrichtung genutzt werden können. Die Erfahrungen des Projekts haben gezeigt, dass die einzelnen Durchgänge des Seminars auch dazu genutzt werden können, diesen Wissenspool systematisch zu erweitern und zu aktualisieren.
- Die technische Infrastruktur muss frühzeitig vor Seminarbeginn bereitgestellt und getestet werden. Auch hier lässt sich der Aufwand für die Bereitstellung und die Betreuung der Infrastruktur für das einzelne Seminar reduzieren, wenn ihre Nutzung auf Nachhaltigkeit angelegt wird. Standardisierung der technischen Ausrüstung der beteiligten Partner ist hierfür ein wichtiges Instrument. Support für die Behebung technischer Probleme muss explizit vorgesehen werden. Dabei können einzelne Aufgaben durchaus auf TeilnehmerInnen im Seminar übertragen werden, um sie so bei der Bewältigung technischer Aufgaben einzubinden.
- Da die unterschiedlichen Systeme der einzelnen Partnerinstitutionen bei der Vergabe von Leistungsnachweisen zu Konflikten führen können,

sollte das Verfahren des Erwerbs von Leistungsnachweisen und der Bewertung von Studienleistungen vor Beginn des Seminars explizit geregelt werden.

4 Fazit

Insgesamt lässt sich feststellen, dass das didaktische Konzept des Seminars sowohl von Studierenden als auch von Lehrenden als interessant und zukunftsfruchtig angesehen wurde. Dabei sind sowohl die Möglichkeit, Formen der verteilten Kooperation zu erproben wie auch die Perspektive, Inhalte in das Studium zu integrieren, die an der eigenen Hochschule nicht angeboten werden, in gleicher Weise attraktiv. Die überwiegende Zahl der Studierenden, die an einem der drei Seminardurchgänge teilgenommen haben, würde wieder ein virtuelles Seminar besuchen. Bemerkenswert ist dabei, dass virtuelle Lehrformen fast ausschließlich als sinnvolle Ergänzung der Präsenzlehre angesehen werden. Der Vision einer vollständig virtuellen Hochschule stehen die Studierenden eher skeptisch gegenüber.

Der letzte Durchgang des Seminars, der im Rahmen des Projekts durchgeführt und evaluiert wurde, war durch seine Rahmenbedingungen den Anforderungen des Regellehrbetriebs stark angenähert. Es zeigt sich, dass das Konzept des virtuellen Seminars für die Integration dieser Veranstaltungsform in den Regellehrbetrieb durchaus tragfähig ist. Eine auf Nachhaltigkeit angelegte Nutzung der in dem Projekt aufgebauten inhaltlichen Ressourcen sowie der technischen und organisatorischen Infrastruktur würde schon nach wenigen Nutzungszyklen die mit dieser Veranstaltungsform verbundene Mehrbelastung für Lehrende und Studierende in einem vertretbaren Rahmen halten.

5 Literatur

- Caroli, Folker et al. (1996). Multimedia. Technologien, Entwicklungen, Chancen und Risiken für das Saarland. Saarbrücken: Arbeitskammer des Saarlandes. (AK Beiträge Jg. 9, H. 5).
- Caroli, Folker (2000). Internetgestützte Seminare. Innovation für eine Form der universitären Lehre. In: Friedrich Scheuermann (Hrsg.). Campus 2000. Lernen in neuen Organisationsformen. Münster: Waxmann. (= Medien in der Wissenschaft 10). S. 149 - 160.
- Caroli, Folker, Oliver Zandner (2000). Seminare im Internet. Ein Werkstattbericht aus dem Projektverbund "Virtueller Campus Hannover – Hildesheim – Osnabrück". In: Stefan Uellner und Volker Wulff (Hrsg.). Vernetztes Lernen mit digitalen Medien. Proceedings der ersten Tagung "Computergestütztes Kooperatives Lernen (D-CSCL 2000)" 23. - 24. März 2000 in Darmstadt. Heidelberg: Physica. S. 183 - 204.

- Caroli, Folker, Christa Hauenschild, Daniela Hellmann (2001). Verteilte Kooperation in virtuellen Seminaren. In: Erich Wagner, Michael Kindt (Hrsg.). Virtueller Campus. Szenarien - Strategien - Studium. Münster: Waxmann 2001. S.483 - 492.
- Gräsel, Cornelia (1997). Problemorientiertes Lernen. Strategie, Anwendung und Gestaltungsmöglichkeiten. Göttingen: Hofgrete.
- van Haaren, Kurt (Hrsg.) (1997). Arbeit im Multimedia-Zeitalter: die Trends der Informationsgesellschaft Hamburg.
- Kerres, Michael (2001). Mediengestützte Lehre und Hochschulstrukturen. In: Erich Wagner, Michael Kindt (Hrsg.). Virtueller Campus. Szenarien - Strategien - Studium. Münster: Waxmann 2001. S.293 - 302.
- Klimsa, Paul (1993). Neue Medien und Weiterbildung: Anwendung und Nutzung in Lernprozessen der Weiterbildung. Weinheim: Dt. Studien-Verlag.
- Lehren und Lernen mit neuen Medien (1999). Lehren und Lernen mit neuen Medien. Plattformen, Modelle Werkzeuge. Beiträge der Fachtagung des Landesarbeitskreises Multimedia des Landes Niedersachsen 25. u. 26. November 1999 in Hildesheim. Hildesheim: Geschäftsstelle Multimedia.
- Mandl, Heinz, Hans Gruber, Alexander Renkl (1997). Situiertes Lernen in multimedialen Lernumgebungen. In: Ludwig J. Issing, Paul Klimsa. Information und Lernen mit Multimedia. 2. Aufl. Weinheim: Beltz. S.167-178.
- Picot, Arnold (Hrsg.) (1997). Telekooperation und virtuelle Unternehmen. Auf dem Weg zu neuen Arbeitsformen. Heidelberg: Decker.
- Reichwald, Ralf et al. (1997). Telekooperation. Verteilte Arbeits- und Organisationsformen. Berlin, Heidelberg: Springer.
- Schmidt, Christiane (2000). Evaluation und Begleitforschung "Lernen im Netz": Forschungsdesign und erste Ergebnisse. In: Gesellschaft für Medien und Wissenschaft (Hrsg.). Virtueller Campus 99. Heute Experiment - morgen Alltag? Münster: Waxmann.
- Schmidt, Christiane (in Vorbereitung). Forschendes Lernen mit Neuen Medien. Eine qualitative Untersuchung zu Lernerfahrungen in internetbasierten Seminaren. Habilitationsschrift Hildesheim.
- Schulmeister, Rolf (1997). Grundlagen hypermedialer Lernsysteme: Theorie – Didaktik – Design München : Oldenburg.
- Uellner, Stefan und Volker Wulf (Hrsg.) (2000). Vernetztes Lernen mit digitalen Medien. Proceedings der ersten Tagung „Computergestütztes Kooperatives Lernen (D-CSCL 2000)“ am 23. - 24. März 2000 in Darmstadt. Heidelberg: Physica.
- Virtueller Campus (2002). Virtueller Campus Sprache - Wissen - Information. Abschlussbericht. Hildesheim (in Vorbereitung)
- Wagner, Daniela, Folker Caroli, Christa Hauenschild, Christa Womser-Hacker (2002). Evaluation: Virtuelles Seminar „Multilinguale Informationssysteme“. Im Rahmen des Verbundprojekts „Virtueller Campus Sprache – Wissen Information“. In: Virtueller Campus Sprache - Wissen - Information. Abschlussbericht. Hildesheim (in Vorbereitung)



Echtzeit-Fahrerbeanspruchungsschätzung

*Walter Piechulla¹, Christoph Mayser², Helmar Gehrke³,
Winfried König³*

¹Universität Regensburg,
Institut für Experimentelle Psychologie,
Universitätsstraße 31,
93053 Regensburg
walter.piechulla@
psychologie.uni-
regensburg.de

²BMW Group,
Forschung, Vorentwicklung, Konzepte - Mensch
Maschine Interaktion,
Max-Diamand-Str. 13,
80788 München
christoph.mayser@
bmw.de

³Robert Bosch GmbH,
FV/SLN Forschung und
Vorausentwicklung,
Postfach 300240,
70442 Stuttgart,
winfried.koenig@
de.bosch.com
helmar.gehrke@
de.bosch.com

1 Einleitung

Zur Anpassung der Mensch-Maschine-Schnittstelle von Personenkraftwagen an situationsspezifische Anforderungen wurde ein System entwickelt, das verschiedene Verkehrssituationen erkennt und hinsichtlich der zu erwartenden Beanspruchungswirkung auf den Fahrer beurteilt¹. Als Beispiel für eine Anpassung der Benutzerschnittstelle dient die Unterdrückung eingehender Telefonanrufe in Situationen, für die eine hohe Fahrerbeanspruchung zu erwarten ist. Diese Idee wurde Anfang der 90er Jahre im Projekt Generic Intelligent Driver Support (Michon, 1993) erstmals erwähnt. Wir stellen ein funktionsfähiges System vor, das wir im realen Straßenverkehr erprobt haben. Die Bean-

¹ Die berichteten Arbeiten gehören zum Projekt SANTOS (Situations-angepasste und Nutzer-Typ-zentrierte Optimierung von Systemen zur Fahrerunterstützung), gefördert mit Mittel des Bundesministeriums für Bildung, Wissenschaft, Forschung und Technologie unter dem Förderkennzeichen 19 S 9826 A/B. Projektleitung: C. Mayser, BMW Group und Dr.-Ing. W. König, Robert Bosch GmbH. Kooperationspartner: Technische Universität Darmstadt, Fachgebiet Fahrzeugtechnik; Technische Universität Darmstadt, Institut für Arbeitswissenschaft; Technische Universität Dresden, Institut für Verkehrswegebau; Technische Universität München, Lehrstuhl für Ergonomie; Universität Regensburg, Lehrstuhl für Psychologie II, Universität Würzburg, Würzburger Institut für Verkehrswissenschaften. Projekt-Homepage: <http://www.santosweb.de>.



spruchungsschätzung beruht dabei einerseits auf einer Situationsklassifikation, die in einer geografischen Datenbank (sogenannte feindigitale Karte FDK) gespeichert ist. Andererseits werden dynamische Ereignisse des Verkehrsgeschehens wie die Annäherung an ein vorausfahrendes Fahrzeug oder starke Verzögerungen als Prädiktoren erhöhter Fahrerbeanspruchung verwendet. Außerdem wird versucht, Umweltbedingungen wie Regen, Reibwert (Griffigkeit der Fahrbahnoberfläche) und den Unterschied zwischen Tag- und Nachtfahrt zu berücksichtigen. Die Einschätzungen von 20 Fremdbeurteilern zeigen, dass es tatsächlich möglich ist, eine sinnvolle Beanspruchungsprognose vorzunehmen.

2 Übersicht über das Vorgehen

Eine 27 km lange Teststrecke wurde zunächst gefilmt, dann anhand der Videoaufzeichnung nach der Taxonomie für Verkehrssituationen von Fastenmeier (1995) klassifiziert. Die Übergänge zwischen den 186 resultierenden Situationen wurden dann vor Ort mit differenziellem GPS (DGPS) vermessen und die Situationsübergänge in die FDK-Datenbank eingetragen. Zur Schätzung der relativen Beanspruchungswirkung der 22 vorkommenden Situationsklassen wurden Beanspruchungsversuche im Sekundäraufgabenparadigma verwendet: In neun Realfahrten mussten die Versuchspersonen eine visuelle Suchaufgabe in einem Lauftext-Display bearbeiten. Als Maß für die visuelle Restverarbeitungskapazität der Fahrer wurde die Anzahl der Blickzuwendungen zur Sekundäraufgabe pro Sekunde verwendet. Das Maß differenziert gut (ANOVA, $p \leq .006$) zwischen leicht zu bewältigenden Situationen (definiert nach Fastenmeier), in denen die Fahrer relativ viele Blicke auf die Sekundäraufgabe richten und schwierigen Situationen, in denen die Fahrer den Blick seltener vom Verkehrsgeschehen abwenden. Die Beanspruchungsindizes (1 - mittlere Blickzuwendungshäufigkeit) für die 22 Situationsklassen werden aus diesem Maß abgeleitet. Das Demonstrator-Fahrzeug, in dem der Beanspruchungsschätzer realisiert wurde, bestimmt ständig per DGPS-Präzisionsortung seinen Standort und errechnet den voraus liegenden Streckenabschnitt (sogenannte „Fahrschlauch-Prognose“). Die Beanspruchungsindizes der Situationen werden dabei gewichtet und zu einem Beanspruchungsschätzwert verrechnet.

3 Implementierung eines belastungsadaptiven Kommunikationssystems

Ein handelsüblicher differenzieller GPS-Empfänger liefert die Positionsbestimmung für den Maptracker (Software), der aus der feindigitalen Karte den Fahrschlauch berechnet. Die Beanspruchungsschätzung verwendet außer Fahrschlauchdaten auch Daten des Adaptive Cruise Control (ACC / ein Abstandsregel-Assistenzsystem), sowie Daten von Fahrzeug-Bussen zur Nachgewichtung des Schätzwertes. Beim Überschreiten eines Schwellwertes werden eingehende Telefonanrufe nicht mehr an den Fahrer signalisiert (Telefonklingeln), sondern in die Mailbox (Anrufbeantworter) umgeleitet.

4 Evaluation

Am Evaluationsversuch nahmen zwölf Versuchspersonen teil, sechs davon waren fahrerfahren, sechs Führerscheinneulinge. Jede Versuchsperson fuhr die Versuchsstrecke dreimal: Einmal ohne Fahrerassistenzsysteme, einmal mit ACC und Heading Control (HC / ein Spurhalte-Assistenzsystem) und einmal mit ACC, HC und dem Beanspruchungsschätzer, der Telefonanrufe abblockt, wenn die Beanspruchungsschätzung hohe Werte annimmt. In jeder Fahrt wurde der Fahrer genau zehn Mal angerufen und am Telefon nach der Lösung einer einfachen Kopfrechenaufgabe gefragt (zweistellige Zufallszahl + 12 = ?). Damit wurde ein Telefongespräch mittleren Schwierigkeitsgrades simuliert. Unter der Bedingung mit aktiviertem Beanspruchungsschätzer waren bis zu 64 Anrufversuche notwendig, da bis zu 54 Anrufe vom System abgewiesen wurden. Die Videosequenzen dieser 30 x 12 Gesprächssituationen wurden später von 20 anderen Versuchspersonen im Blindversuch auf einer sechsstufigen Ratingskala danach eingeordnet, wie stark ihrem Eindruck nach die zusätzliche Beanspruchung durch das Telefongespräch war. Das selektive Abblocken von Telefonanrufen führt für die erfahrenen Fahrer zu einer Beanspruchungsreduktion (Wilcoxon Signed-Rank Test, $p \leq .0005$). Dies kann nur auf das Verlagern der Telefongespräche in gering beanspruchende Situationen zurückgeführt werden. Für die Fahranfänger ist kein Unterschied zwischen den Bedingungen nachweisbar.

5 Schlussfolgerungen

Eine experimentelle feindigitale Karte und DGPS-Präzisionsortung ermöglichen es, Fahrzeuge mit einer Art "Bewusstsein" der Verkehrsumgebung auszustatten. Das System ist zwar noch auf ein kleines Testgebiet beschränkt,

steht aber mit der Einführung genauerer digitaler Karten in nicht allzu ferner Zukunft möglicherweise flächendeckend zur Verfügung. Messdaten intelligenter Sensoren wie der Abstand zum vorausfahrenden Fahrzeug, aber auch relativ einfach zu bestimmende Größen wie die Längsbeschleunigung können wertvolle Hinweise auf die situationscharakteristische mentale Beanspruchung des Fahrers liefern. Weiterentwicklungen des vorgestellten Systems könnten in Zukunft Fahraufgabe und Kommunikation komfortabler und sicherer machen.

6 Literatur

- Fastenmeier, W. (Hrsg.) (1995). Autofahrer und Verkehrssituation. Neue Wege zur Bewertung von Sicherheit und Zuverlässigkeit moderner Straßenverkehrssysteme. Köln: Verlag TÜV Rheinland.
- König, W., Weiß, K.E., Gehrke, H. und Haller, R. (2000). S.A.N.T.O.S Situationsangepasste und Nutzer-Typ-zentrierte Optimierung von Systemen zur Fahrerunterstützung. In H. Bubb (Hrsg.) Ergonomie und Verkehrssicherheit. Dokumentation der Herbstkonferenz der GfA in München, 12.-13.10.2000 (S. 107-113). München: Utz.
- Michon, J.A. (Hrsg.) (1993). Generic intelligent driver support. A comprehensive report on GIDS. London: Taylor & Francis.
- Rockwell, T.H. (1988). Spare visual capacity in driving revisited: new empirical results for an old idea. In A. G. Gale et al. (Hrsg.), Vision in Vehicles II. (S. 317-324), Amsterdam: Elsevier.



„Wörter des Tages“ - Tagesaktuelle wissensbasierte Analyse und Visualisierung von Zeitungen und Newsdiensten

Uwe Quasthoff¹, Matthias Richter¹, Christian Wolff²

¹Universität Leipzig,
Institut für Informatik,
Augustusplatz 10/11, 04109 Leipzig
quasthoff@informatik.uni-leipzig.de,
matthias@vielfalt.de

²Universität Regensburg
Institut für Medien-, Informations- und
Kulturwissenschaften
93040 Regensburg
christian.wolff@sprachlit.uni-
regensburg.de

Abstract

In this paper we describe a web service which presents media analysis results as both, structured lists of *Words of the Day*, as well as visualizations of relevant concepts in their semantic context and concept relevance timelines. Media analysis is based on a daily collection of online newspapers and news services. For each day, this collection is processed by a text mining suite.

1 Einführung

Dieser Beitrag versucht aufzuzeigen, wie Text Mining-Verfahren für die Analyse von Medienprodukten genutzt werden können und so als „angewandte Medieninformatik“ einen interdisziplinären Beitrag zur Medienanalyse leisten können. Es wird ein im World Wide Web verfügbarer Informationsdienst vorgestellt, der tagesaktuell überregionale Online-Medien auswertet und begriffsbasiert die jeweils als relevant erkannten Konzepte als „Wörter des Tages“ präsentiert. Die „Wörter des Tages“ wurden im Rahmen des Projekts „Deutscher Wortschatz“ am Institut für Informatik der Universität Leipzig entwickelt (vgl. Quasthoff & Wolff 2000) und sind im World Wide Web unter <http://www.wortschatz.uni-leipzig.de/wort-des-tages/> verfügbar.

2 Tagesaktuelle Analyse von Online-Informationsdiensten

Die Analyse von Online-Texten umfasst folgende Schritte:

- Quellenauswahl und –erfassung



- Quellenanalyse durch Text Mining
- Begriffsselektion (Kandidaten für „Wörter des Tages“)
- Kategorisierung und Überarbeitung von Kandidatenlisten

Die Quellenauswahl erfolgt unter Heranziehung technischer, quantitativer und qualitativer Merkmale. Dabei werden vor allem die Online-Sites überregionaler Medien (z. B. www.sueddeutsche.de, www.spiegel.de). Zum derzeitigen Stand der Quellengrundlage umfasst ein solches Tagescorpus ca. 20.000 Sätze. Die jeweils über Nacht gesammelten Quellen werden einer Text Mining-Analyse durch die im Projekt „Deutscher Wortschatz“ entwickelten Werkzeuge unterzogen (vgl. Heyer, Quasthoff & Wolff 2000; Quasthoff & Wolff 2002). Diese besteht aus folgenden Schritten:

0. Bereitstellung eines um den Faktor 1000 umfangreicheren Referenzcorpus.
 1. Textsegmentierung, insbesondere Satz- und Wortsegmentierung.
 2. Indexierung der Beispieltex-te und quantitative Erfassung der Einzelbegriffe.
 3. Berechnung relevanter Satz- und Nachbarschaftskollokationen.
 4. Speicherung der Ergebnisse in einer relationalen (Tages-)Datenbank.

Nach Durchführung der Text Mining-Analyse für das jeweilige Tagescorpus muss eine „handhabbare“ Menge von Wörtern des Tages selektiert werden, die sich für eine Präsentation im Rahmen eines Web Service eignet. Da sich die Eignung von Wörtern als Aktualitätsindikator der Medienanalyse nicht nach einem einzelnen Parameter richten kann, stehen für die Auswahl der „Wörter des Tages“ die folgenden drei statistischen Basisparameter zur Verfügung: Frequenz im aktuellen Tagescorpus, Frequenz im Referenzcorpus (Deutscher Wortschatz) und relative Übergewichtung eines aktuellen Begriffs.

3 Ergebnisaufbereitung und Visualisierung

Die Ausgabe der Wörter des Tages lässt sich in folgende Bereiche gliedern:

- Ausgabe der Wörter des Tages (nach Kategorien geordnet)
- Ausgabe der Belegstellensammlung zu einem Wort des Tages
- Ausgabe eines Assoziationsgraphen, der Beziehungen eines Wortes des Tages zu anderen im Tagescorpus enthaltenen Konzepten darstellt
- Graphische Ausgabe des „Aktualitätsverlaufs“ von Worten des Tages

Die Startseite des *web service* gibt eine Übersicht zu den aktuellen Wörtern des Tages, die nach den ihnen zugeordneten Kategorien sortiert ausgegeben werden (Abb. 1). Bei der Analyse werden Belegstellen für die verschiedenen

„Wörter des Tages“ in einer Datenbank gesammelt und können online abgefragt werden; auch der Zugriff auf die Quelldokumente ist möglich.

Wortschatz : Wörter des Tages : 20.06.2002	
Sportler, Trainer, Funktionäre	Frank Baumann · Fritz Walter · Guus Hiddink · Hiddink · Jung-Hwan · Kloiber · Moreno · Neuville · Ronaldo · Rudi Völler · Völler
Sport	Bayer Leverkusen · Eintracht · Eintracht Frankfurt · Fifa · Formel 1 · Fußball · Fußball-WM · Golden Goal
Politiker	Bundeskanzler Gerhard Schröder · Clement · Eichel · Hamid Karsai · Künast · Stoiber
Organisation	AMD · Apple · Creditreform · Deutsche Post · Deutsche Telekom · Infineon · Kirch Media · KirchMedia · MLP · Oracle · Taliban · WCM
Ereignis	Gewitter · Insolvenz · Kieler Woche · Selbstmordanschlag · Sommerferien
Schlagwort	Achtelfinale · Anschläge · Arbeitskampf · Attentäter · Audiogalaxy · Bildungspolitik · Bürgschaft · Hitze · Loja Dschirga · Palästinenser · Rentenreform · Riester-Rente · Selbstmordattentäter · Spirit · Studiengebühren · T-Aktie · Terrorismus · Viertelfinale · Zuwanderung · Zuwanderungsgesetz
Ort	Afghanistan · Jerusalem · Kabul · Korea · NRW · Nürburgring · Paraguay · Senegal · Seogwipo · Südkorea · Wuppertal
Personen aus Kunst, Kultur und Wissenschaft	
sonstige Personen	Duri Lozza · Koreaner · Späth · Südkoreaner
«19.06.2002» Wörter des Tages	

Abbildung 1: Webbasierte Ausgabe der *Wörter des Tages* vom 20. Juni 2002

Neben der textbasierten Präsentation einer jeweils aktuellen Auswahl von „Wörtern des Tages“ lassen die zugrundeliegenden Text Mining-Verfahren auch eine *Visualisierung begrifflicher Zusammenhänge* bzw. die Darstellung des *zeitlichen Verlaufs der Aktualität* von Wörtern des Tages zu. Abb. 2 zeigt für das derzeit häufig als „Wort des Tages“ vertretene Konzept *T-Aktie* den tagesaktuellen Assoziationsgraphen vom 19. Juni 2002 sowie den Assoziationsgraphen aus dem Referenzcorpus. Dabei wird deutlich, dass die den Tagescorpora entnommenen Graphen jeweils aktuelle Relationen hervorheben (z. B. *Aktienoptionen*), während im Referenzgraphen eher grundsätzliche Beziehungen deutlich werden (*Volksaktie*, *Dax*, *Ausgabekurs*, *Ron Sommer*).

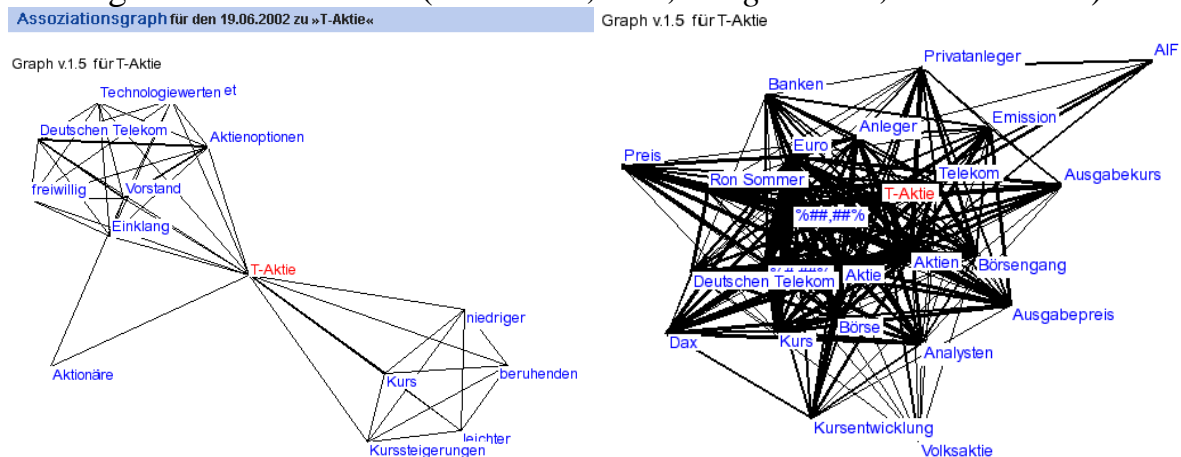


Abbildung 2: Graphen für *T-Aktie* aus Tagescorpus (links) bzw. Referenzcorpus (rechts)

Es liegt nahe, die tagesbezogenen Ergebnisse zu den Wörtern des Tages auch als längerfristige Entwicklungen zu untersuchen. Da die Analysecorpora für jeden Tag gesondert verfügbar sind, lassen sich die Auswahlmengen unmittelbar miteinander vergleichen. Die Darstellung erfolgt mit Hilfe eines Liniendiagramms, das sich für die Darstellung von Verlaufsentwicklungen sehr gut

eignet. Dabei werden als Graphen jeweils die relativen Aktualitätswerte des ausgesuchten Begriffs sowie seiner jeweils stärksten Kollokationen angezeigt. Zusätzlich werden diejenigen Tage markiert, an denen der jeweilige Begriff unter den Wörtern des Tages war. In Abb. 3 wird der Aktualitätsverlauf zu *Michael Schumacher* dargestellt. Deutlich erkennbar ist dabei der Anstieg im 14-Tages-Rhythmus (Formel 1-Kalender).

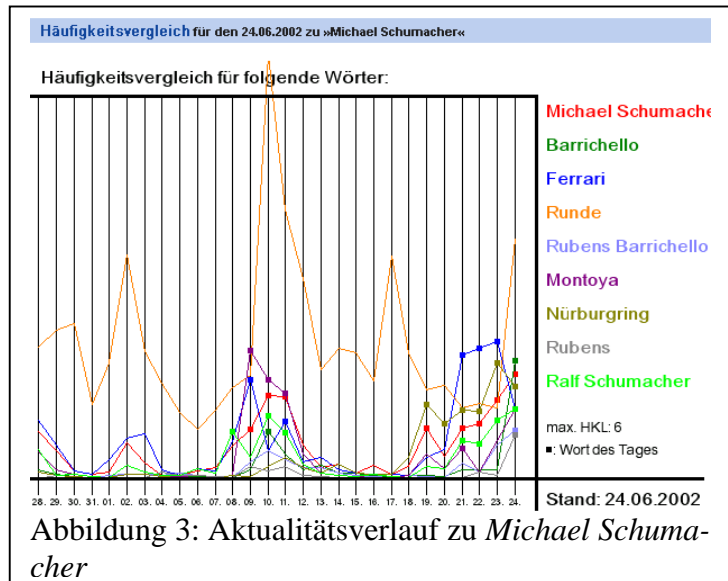


Abbildung 3: Aktualitätsverlauf zu *Michael Schumacher*

4 Fazit: Bewertung und Nutzung

Abschließend stellt sich die Frage nach einer Bewertung der jeweils getroffenen Auswahl von Ereignissen und Personen. Selbst bei exakter Definition von Randparametern der Analyse wie Medienauswahl, Zeit, Ort, gesellschaftlich-politisches Bezugssystem oder Zusammensetzung und Anzahl der Rezipienten ist die Definition *objektivierbarer Kriterien* für die Relevanzmessung von Personen und Ereignissen problematisch. Insofern ist auch der hier beschriebene Ansatz, durch quantitative Medienanalyse Aussagen über die zeitbezogene Relevanz zu treffen, nur eine Annäherung. Erstrebenswert ist die Entwicklung eines *media impact index*, der eine Bewertung von Ergebnissen, wie sie hier beschrieben sind, zulässt (vgl. dazu Posner 2001).

5 Literatur

Quasthoff, Uwe; Richter, Matthias; Wolff, Christian (2002). Wörter des Tages - medienanalytische Motivation, texttechnologische Konzeption und Realisierung als Web Service. Technical Paper, Universität Leipzig, Institut für Informatik, Juli 2002.

Quasthoff, Uwe; Wolff, Christian; "An Infrastructure for Corpus-Based Monolingual Dictionaries". In: Proc. LREC-2000. Second International Conference on Language Resources and Evaluation. Athens, May / June 2000, Vol. I, 241-246.

Posner, Richard A. (2001). Public Intellectuals. Cambridge/MA.: Harvard University Press.



Mehr oder weniger Mobilität durch Virtualität?

Wissens- und Informationsmanagement über interaktive Netzwerke

Steffen H. Elsner

Bahnhofstraße 32a
D-94099 Ruhstorf/Rott
Fax: +49 (0) 85 03 - 85 75
elsner@wissdok.net

Zusammenfassung

Am Beispiel von Knowledge Communities (Wissensgemeinschaften) wird problematisiert, ob - und in welcher Weise - Virtualität mehr oder minder Mobilität bedeutet.

Abstract

By the example of Knowledge Communities (Wissensgemeinschaften) it will be discussed, whether - and in which particular way - virtuality means more or less mobility.

Obgleich sich die Ursprünge des 'Netzwerk-Phänomens' weit zurückverfolgen lassen, so haben Netzwerk-Konzepte, ausgelöst durch die geradezu revolutionären Entwicklungen im Bereich der Informations- und Kommunikationstechnologie, in jüngster Zeit einen bemerkenswerten Aufschwung erfahren. Keine Branche vermag sich heutzutage auf Dauer erfolgreich der gezielten Einführung und Förderung netzartiger Organisationsformen und Beziehungsgeflechte zu entziehen.

Dabei taucht das Konzept des Netzwerks in der Managementliteratur zum ersten Mal in den siebziger Jahren auf. Inspiriert durch das Aufkommen von lokalen Computernetzen (LAN) und die Neuerungen im weiteren Bereich der Kommunikationstechnologien entsteht ein neues Paradigma: „Das Unternehmen als kommunizierende Organisation.“¹ Netzwerke dienen der Verbreitung von Informationen und ermöglichen Teamwork an verschiedenen Standorten. Netzwerkstrukturen setzen sich fortan mehr und mehr durch und transformie-

¹ Sveiby/Lloyd 1990: S.63.



ren die ehemals klar strukturierten und vergleichsweise leicht zu managenden Organisationseinheiten zu vielfach verwobenen Netzwerkgebilden mit unscharfen Grenzen.

Im Vergleich zu traditionellen Organisationsstrukturen früherer Zeiten unterliegen (virtuelle) Netzwerke heutzutage - jedenfalls vom Prinzip her - keinen geographischen und raum-zeitlichen Begrenzungen. Virtualität erweitert ihren Wirkungskreis auf immer mehr Menschen, institutionelle bzw. organisationale Einheiten und geographische Räume. Insofern kann auch das Potential des Internets zur Herstellung, Aufrechterhaltung und Festigung von (virtuellen) Beziehungen (sog. 'Online Relations') schwerlich überschätzt werden. Zudem entwickeln sich darauf aufbauend neue Arbeitsformen wie etwa *eWorking* oder Telearbeit bzw. die sog. Homeoffice-Lösungen.

Obgleich sie aufgrund dieser Entwicklungen etwas in den Hintergrund geraten sind, haben persönliche Kontakte außerhalb der virtuellen Netzwerke nichts von ihrer Wichtigkeit eingebüßt. Der zwischenmenschliche Kontakt lässt sich durch Computernetze schlechterdings nicht substituieren. Doch können diese zur Stärkung bestehender Beziehungen ebenso beitragen, wie sie es ermöglichen, sehr viel umfangreichere Netzwerke zu schaffen und zu unterhalten, als dies bisher möglich war. Von daher gesehen verdienen Netzwerke, sowohl virtuelle als auch persönliche, die vollste Aufmerksamkeit und den ihnen gebührenden Stellenwert nicht nur innerhalb organisationstheoretischer Erörterungen, sondern gerade auch im realen Unternehmensalltag.

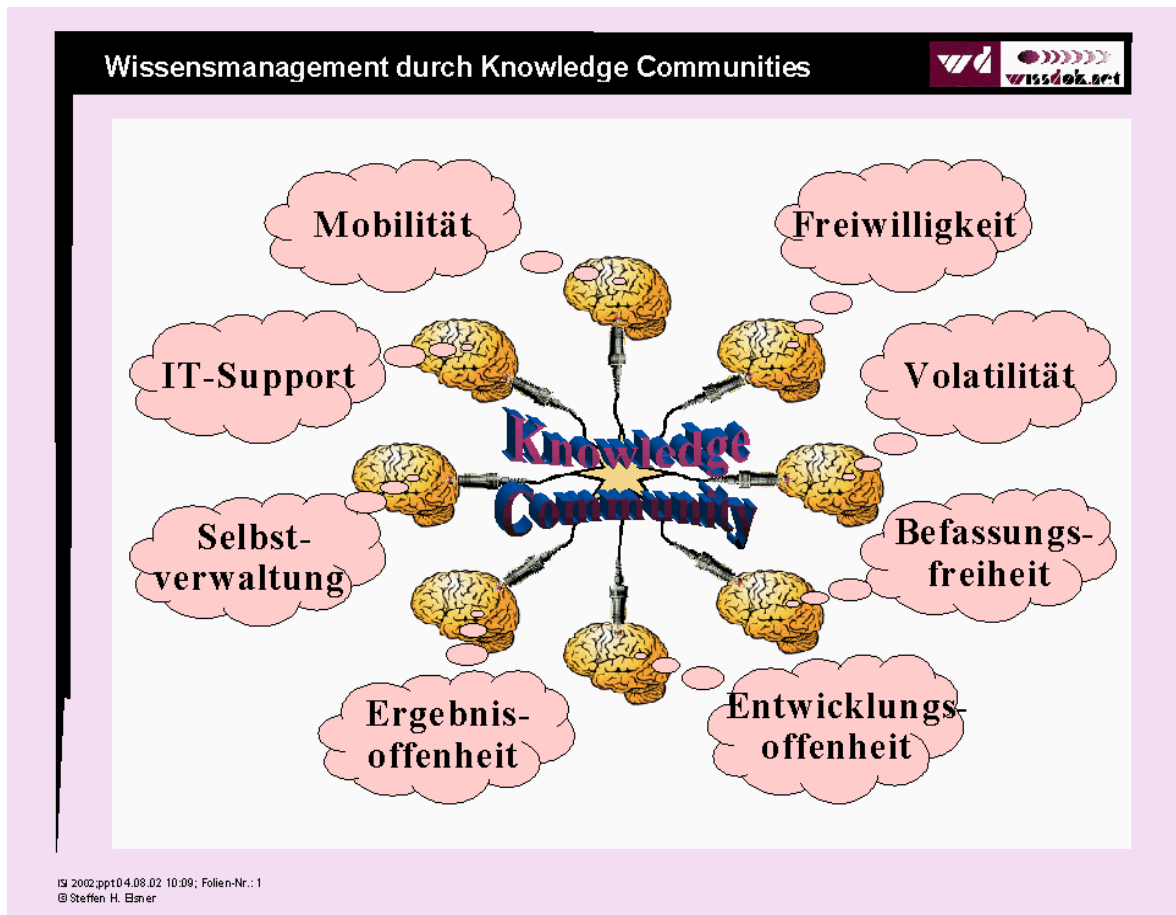
Das gilt für den Bereich des Wissens- und Informationsmanagements sicherlich im besonderen Maße: Von der zugrundeliegenden Idee her funktionieren die Wissensorganisation und der Wissensaustausch primär durch formelle oder informelle Wissensgemeinschaften, auch als 'Knowledge Communities' bezeichnet. North et al. verstehen darunter „... über einen längeren Zeitraum bestehende Personengruppen, die Interesse an einem gemeinsamen Thema haben und Wissen gemeinsam aufbauen und austauschen wollen. Die Teilnahme ist freiwillig und persönlich. Wissensgemeinschaften sind um spezifische Inhalte gruppiert.“²

Ein spezifisches Ziel des ganzheitlich ausgerichteten Knowledge Managements muss demzufolge im Aufbau und in der Kultivierung von vernetzten Wissensgemeinschaften zum (virtuellen) Austausch von Informationen, Wissen, Meinungen und Know-how bestehen.³ Denn schließlich hat das Know-

² North/Romhardt/Probst 2000: S. 54.

³ Vgl. hierzu etwa Elsner 2002, insbesondere S.270ff.

ledge Management zur maßgeblichen Aufgabe, 'den Wissensschatz' einer Organisation transparent(er) und damit möglichst breit zugänglich und nutzbar zu machen.



Für einen prosperierenden Wissensaustausch ist eine effektive Kommunikationsinfrastruktur unerlässlich. IT nimmt hierbei die Rolle eines Hilfsmittels ein, um Informationen und explizierte Wissensbestände einspeichern und zwischenspeichern zu können sowie umfassende Kommunikationsmöglichkeiten zu eröffnen. Das Unterstützungspotential innovativer Technologien für den Zugriff auf und die Verteilung von Wissen ist unbestritten – insofern lässt sich räumliche Mobilität von Mitarbeitern zumindest partiell vermeiden. Gleichwohl können IT-gestützte Informationssysteme den direkten interpersonellen Wissens- und Erfahrungsaustausch niemals vollständig ersetzen. Letztlich gilt, dass die adäquate Kommunikationsform maßgeblich von der Art des auszutauschenden Wissens abhängig ist. Als grobe 'Faustregel' kann hier gelten: Je mehr implizite Wissensbestandteile ausgetauscht werden sollen, desto stärker ist auf den Face-to-Face-Kontakt zu setzen. Je mehr explizites Wissen ausgetauscht wird, desto mehr kann auf virtuelle Kommunikationsstrukturen zurückgegriffen werden.

Literaturverzeichnis

- Elsner, Steffen H.*, 2002: »Brain gain statt Brain drain!« – Wissensmanagement durch Knowledge Communities, S. 265-282 in: Ralph Schmidt (Hg.), Content in Context – Perspektiven der Informationsdienstleistung. 24. Online-Tagung der DGI, Frankfurt a.M. 3.-5. Juni 2002. Frankfurt a.M.: Deutsche Gesellschaft für Informationswissenschaft und Informationspraxis.
- North, Klaus / Romhardt, Kai / Probst, Gilbert*, 2000: Wissensgemeinschaften – Keimzellen lebendigen Wissensmanagements, in: IO-Management, Nr. 7/8 2000, S. 52-62 sowie – in leicht modifizierter Form (Vorabversion) – unter Url: <http://www.cck.uni-kl.de/wmk> (04.08.2002).
- Sveiby, Karl Erik / Lloyd, Tom*, 1990: Das Management des Know-how. Führung von Beratungs-, Kreativ- und Wissensunternehmen. Frankfurt a.M. / New York: Campus.



Mobile Home Security with GPRS

Maurice Danaher, Duy Nguyen

Edith Cowan University
Perth, Australia
m.danaher@ecu.edu.au

Abstract

This paper presents research-in-progress on the development of a home security system for use on mobile devices. Mobile Home Security (MHS) is a system that uses the wireless high-speed technology GPRS. An effective system requires the delivery of good quality images and video to the mobile device. Existing wireless networks have low bandwidth and hence are not capable of providing this service. The introduction of GPRS offers a number of advantages including high bandwidth and low operational cost. Our research is investigating how effectively video and images can be captured and sent across a radio network. Results of this work will provide an awareness of the problems and we will provide guidelines for implementing a mobile security system. Further our research will provide insights into future bandwidth hungry video data applications, not only in wireless security applications but also in other related areas of mobile communication technologies.

1 Introduction

The use of mobile devices that allow users to roam freely without being tied to ground based wires is increasing at a dramatic rate. With the introduction of the Wireless Application Protocol (WAP) for mobile devices we have seen an increasing demand for higher speed mobile applications. WAP enables the mobile user to connect to various Internet applications wirelessly with the ability to roam freely. However the low data speed and cost involved with each connection presents many limitations in the implementation of those types of applications.

For our MHS and other applications involving video data transmission a higher speed mobile service needs to be used to cope with the larger band-



width requirements. GPRS is a relatively new technological advance in the mobile communication market. GPRS enables higher speed data rates and the always-online technology only charges the user for data transfer. Rysavy [1] states that, in theory, speeds of 170 kbps are possible. Operational speeds were in the range of 28 - 40 kbps in early 2001 and in the range 56 - 110 kbps by early 2002. There are however known problems and complications involved with GPRS and the integration of this new technology with the existing radio network [1,2]. Our research aim is not only to create a home security application but also to determine if the system is reliable and effective over the GPRS network. Reliability tests will be performed with respect to the packet switch network over the Global System for Mobile communications network (GSM).

2 The System

Figure 1 shows the major components of the system and the interconnection between the devices. Currently the system consists of a standard generic web cam connected to a personal computer. The main purpose of the computer is to act as a dedicated web cam server. The server needs to have a permanent connection to the Internet with a static IP address. The web cam we are using is a Logitech WebCam Express with USB connection and is capable of capturing video feeds with audio. Still snapshot images can also be captured to produce higher resolution images. As the web cam does not have motion sensors we are using Inetcam iVista™ [3] software to detect movement or changes. The software works by comparing frame after frame for differences, including light intensity differences. Inetcam iVista™ can be configured for different levels of motion sensing and light settings.

All video and image files captured are placed in a local folder on the server which can only be accessed via a File Transfer Protocol (FTP) client. Our server's FTP server allows access only to users with login accounts. A HP Jornada Pocket PC was the chosen receiving device for video and images from our server. This runs WinCE and Pocket Internet Explorer. The GPRS Internet connection will be accessed with an external mobile phone with GPRS compatibility. The mobile phone will act solely as a GPRS modem unit. Our system, however, currently accesses the Internet via a docking station which shares any desktop computer's internet and network resources.

We are about to begin the testing phase of our security system. Tests for reliability of the packet-switch network in comparison to that of the existing circuit-switch will be examined. The main focus of reliability testing is on data

travel from the web cam to the mobile device and vice versa. The data will be presented in the format of compressed images and video feeds with and without audio. Park's [4] prototype of the GPRS architecture has been redesigned to accommodate our test operation. This prototype, which acts as a base to conduct performance calculations, is a simplified version of an actual GPRS network.

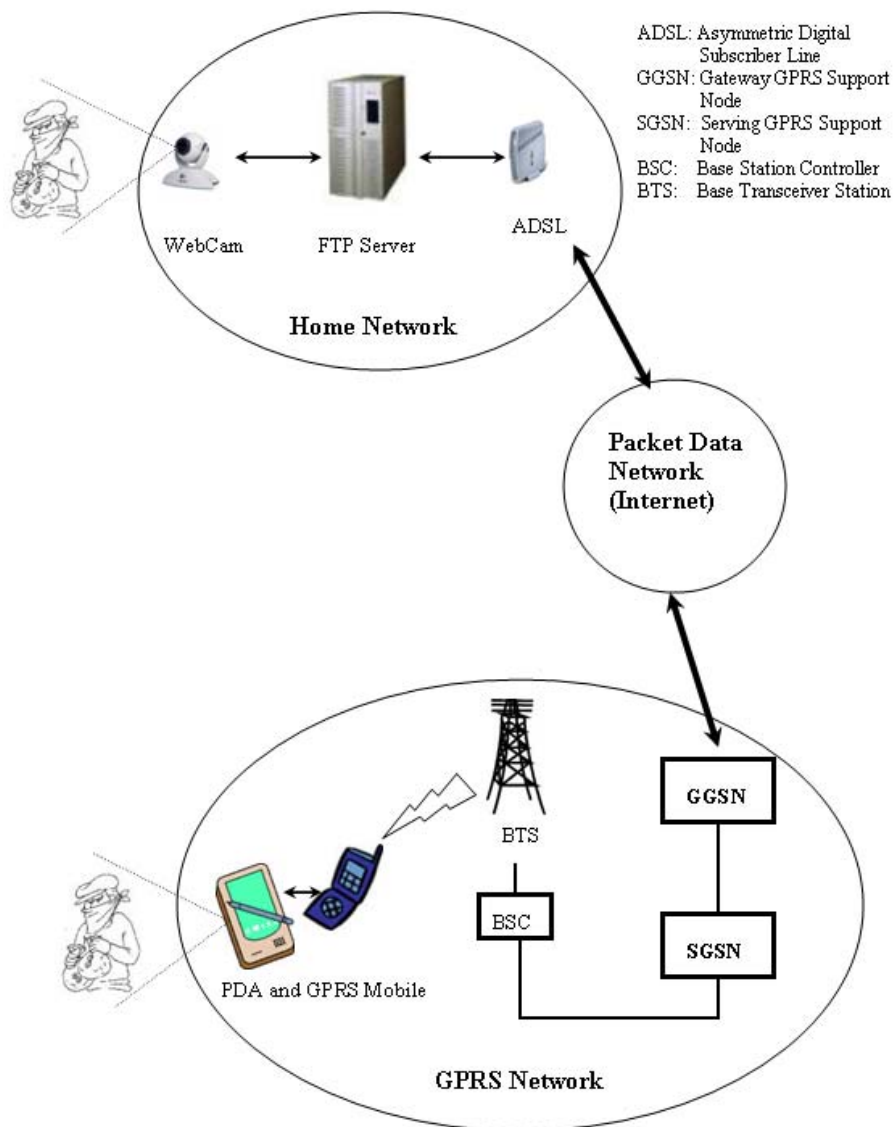


Figure 1 Mobile Home Security

Experiments using different video quality capture rates will be conducted to determine the optimum data stream for the GPRS network. Further, we will investigate potential problems or limitations with the system due to image quality.

3 References

- [1] Rysavy, P., Emerging Technology: Clear Signals for General Packet Radio Service. *Network Magazine*. Dec 2000.
- [2] Vriendt, J. D. et al., Mobile Network Evolution: A Revolution on the Move., *IEEE Communications*, 104- 111, April, 2002
- [3] Inetcam Inc, “iVista Software”, URL: www.inetcam.com, 2002.
- [4] Park, J.-H., Wireless Internet Access for Mobile Subscribers Based on the GPRS/UMTS Network. *IEEE Communications*: 38-49, 2002



Creating a 3D University website with VRML

Maurice Danaher, Chi Wai Lee

Edith Cowan University
Perth, Australia
m.danaher@ecu.edu.au

Abstract

One of the limitations of HTML websites is that even with high quality and aesthetically pleasing graphics, the visitor still experiences a flat 2D world. A three-dimensional world within a website can provide the visitor with a rich, stimulating and interesting experience. Here we describe how we created a model of our university within a website using VRML. The modeling was carried out using 3DstudioMax and video and still images were mapped onto the models. A number of difficulties had to be overcome in order to successfully create the 3D website.

1 Introduction

Up to this point in time websites are designed and developed using 2D graphics. Recently considerable research has been conducted on the development of 3D virtual worlds. Virtual Reality Modeling Language (VRML) is the most widely accepted approach for creating virtual worlds on the Internet. The aim of our work was to create a completely functional 3D website using VRML.

VRML is the new standard file format for describing interactive 3D objects and worlds [1]. We used the final release version of VRML 97 for our research. For viewing the 3D website we chose Cosmo Player [2], a plug-in, by Silicon Graphics. We considered this to be the most fully featured and capable viewing system currently available. Other viewer systems render and navigate in various different ways.

All modelling for this work was performed with 3D StudioMax. This popular software package supports a number 3D file formats. It provides the capability of saving files in a format readable by VRML 97. Figure 1 shows an ex-



ample of a VRML model created in 3D StudioMax. VRML allows us to map images and movies [3] onto the 3D models. One of the objects created in our virtual world is a representation of a television. The purpose of the television is to show visitors various news announcements and other information. This news information was recorded on a digital video camera and after editing it was mapped onto the television.



Figure 1: Example of VRML file created in 3D studio Max

The computing equipment used in this work was a Pentium 4, 1.7 GHz with 256 MB RAM and running Windows XP Professional. Our video equipment consisted of a Sony digital video camera and an EZDV video editing card.

2 Method

Here we give an overview of the method used for creating the website. The process is shown in diagrammatic form in Figure 2. There are two ways to place a VRML file into a web page. The first option is to create a link to the VRML file. For example, “` Enter my virtual world `”. This option allows us to maximize the size of the VR world and provides complete control of the VRML scenes, but the scene will appear in either an external window or in a separate web page from the rest of the document. The second option is to place a VRML scene within a page using a plug in. For example: “`<embed src=“VRTV.wrl” height=100%, weight=100%>`”. This is a standard embed tag which allows us to adjust the scene size in pixels or percentages.

To display a 2D movie on a 3D object the movie is mapped onto the object. In our case the movie is mapped onto the 3D television object. The movie was recorded on the digital video camera and then transferred to the computer. The video sequences were edited on the EZDV video editing board. The video was originally saved in AVI format and later converted to MPEG, a format compatible with VRML. The conversion was done with AVI2MPG.

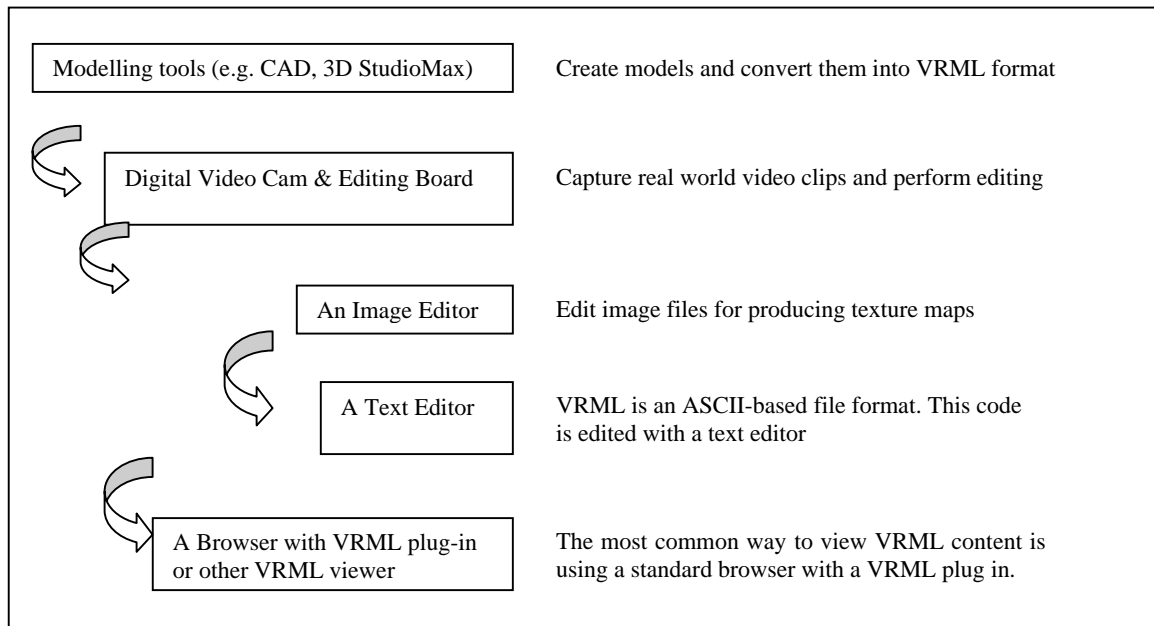


Figure 2: The process to create the 3D VRML website

3 Conclusion

In creating our 3D website a number of problems had to be resolved in the process. Following are some points that should be noted by developers that wish to follow our approach. Particular attention needs to be paid to how the models are created in 3D StudioMax. It is important to watch the polygon count as too many polygons can cause system failure. We recommend that scenes contain not more than 3000 polygons. Also faces that are never seen should be hidden or deleted. Animation that uses simple transforms tends to produce smaller VRML files than animation that uses coordinate interpolation. With coordinate interpolation, the VRML exporter has to calculate each vertex separately, while with simple transforms it calculates for the object as a whole. To control the viewpoint in the virtual world cameras should be placed in appropriate positions, otherwise viewing will default to perspective views. Only standard and multi/sub-object materials should be used, as 3D StudioMax will not export other types of materials.

A standard AVI file is restricted to a maximum size of 2 Gigabytes. This restriction is overcome on the EZDV video editing board with *Reference* technology. Reference AVI files are small AVI files that *point* to other files that contain the actual video data. To facilitate working with these large video files it is best to have a large amount of memory. Also, we could obtain more impressive results with the display of the movie within the website by using a

more powerful processor, more RAM and a hardware accelerated graphics card. We found that Cosmo Player is the best VRML viewer as it can provide full control of the navigation within our VR world. Browsers that are incompatible with Cosmo will cause some features to be disabled or malfunction. The browser we used was Internet Explorer 6.

4 References

- [1] Web 3D Consortium, “Web3D Specifications: VRML 97,”
URL: http://www.vrml.org/fs_specifications.htm
- [2] Cosmo Software, “Cosmo Player”, URL: <http://www.cai.com/cosmo/>
- [3] Walsh, A. E. & Bourges-Sevenier, M, “Core Web 3D,” 1st Edition, Prentice Hall, 2000



Ein prototypisches Werkzeug zur flexiblen Visualisierung komplexer Datenbestände

Thomas Mandl¹, Maximilian Eibl²

¹Informationswissenschaft
Universität Hildesheim
Marienburger Platz 22
31141 Hildesheim
mandl@uni-hildesheim.de

²IZ Sozialwissenschaften
GESIS-Außenstelle
Schiffbauerdamm 19
10117 Berlin
eibl@berlin.iz-soz.de

Visualisierung gilt als eine wichtige Technik, um Benutzern von Informationssystemen die Interaktion mit großen und komplexen Datenmengen zu erleichtern. Insbesondere niedrigdimensionale Dokumentkarten gewannen nicht zuletzt durch erfolgreiche Implementierungen im Internet an Bedeutung. Diese Karten gruppieren hochdimensionale Ausgangsdaten wie beispielsweise Textdokumente auf einer zweidimensionalen Fläche oder in einem dreidimensionalen Raum, wobei die Abstände zwischen den Dokumenten semantische Ähnlichkeit ausdrücken. So entstehen räumliche Anhäufungen von Daten, die inhaltlich zusammengehören: Es werden Thematiken visualisiert.

Trotz zahlreicher Verfahren zur Dimensionalitätsreduktion wie auch zur graphischen Umsetzung der Karten gibt es bislang keine umfassende systematische Untersuchung über die Zweckmäßigkeit der einzelnen Verfahren oder gar der Karten an sich. Der hier vorgestellte Prototyp ist Teil einer solchen momentan durchgeführten Untersuchung. In einem ersten Schritt untersuchen Mandl & Eibl 2001a & b und Eibl & Mandl 2002 die unterschiedlichen Reduktionsmechanismen auf mathematischer Basis unter Verwendung verschiedener Nutzerhypothesen, wobei sich starke Divergenzen herausstellten. In einem zweiten Schritt wird ein flexibel parametrisierbares Werkzeug zur Erstellung unterschiedlichster niedrigdimensionaler Karten erstellt, welches die systematische Analyse der graphischen Umsetzungsmöglichkeiten von Dokumentkarten ermöglicht. Ein erster Prototyp dieses Werkzeugs wird hier vorgestellt. Für die Zukunft sind weitere Schritte wie der inhaltliche Abgleich automatisch erstellter Dokumentkarten mit intellektuell erstellten Themendokumentationen sowie intensive Benutzertests geplant. Abbildung 1 zeigt einige der mit dem Prototyp erstellbaren Formen von Visualisierung.



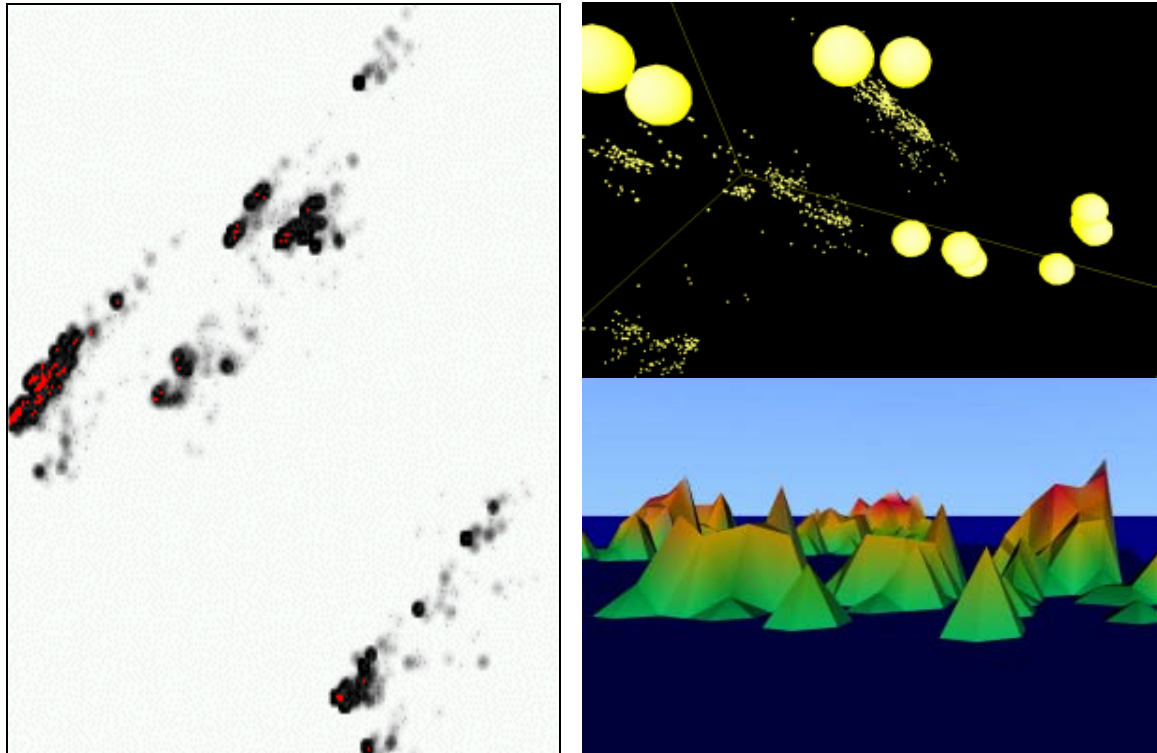


Abbildung 1: Unterschiedliche Visualisierungen eines Datensatzes

In einer ersten Implementierung des Prototyps werden zu einem Satz dimensionalitätsreduzierter Daten verschiedene graphische Darstellungen angeboten, die sich in Dimensionalität, Codierung und Metaphorik voneinander unterscheiden. So werden hier zwei Analysen ermöglicht: Zum einen können anhand der Graphiken die Vor- und Nachteile verschiedener graphischer Umsetzungen systematisch analysiert werden. Zum anderen ermöglichen sie auch bereits eine erste Analyse der zugrundeliegenden Daten, indem Dokumentcluster (Thematiken) und Anomalien (Einzeldokumente) angezeigt werden. Ein weiterer Ausbau des Prototyps soll die genauere interaktive Analyse der Darstellungen erlauben.

Literatur

- Eibl, Maximilian; Mandl, Thomas (2002): Including User Strategies in the Evaluation of Graphic Design Interfaces for Browsing Documents. In: Journal of WSCG. Special issue. Vol. 10, No.1, S. 163-169. http://wscg.zcu.cz/wscg2002/Papers_2002/B89.pdf
- Mandl, Thomas; Eibl, Maximilian (2001a). Evaluating Visualizations: A Method for Comparing 2D Maps. In: Proceedings of the 9th HCI International, New Orleans, 5.-10. August 2001, Band 1, S. 1145-1149.
- Mandl, Thomas; Eibl, Maximilian (2001b). Topographic Maps Based on Kohonen Self Organizing Maps: An Empirical Approach. In: Proceedings of the EUNITE, 13.-14. Dezember 2001, Puerto de la Cruz (Teneriffa), S.467-471.



SELIM: Softwareergonomie für Lernsysteme mit verschiedenartiger lerntheoretischer Basis

Monika Schudnagis, Christa Womser-Hacker

Informationswissenschaft, Universität Hildesheim
Marienburger Platz 22
31141 Hildesheim
{schudnag,womser}@rz.uni-hildesheim.de

1 Einleitung

Im Projekt SELIM¹ (s. Beitrag von KAMENTZ & SCHUDNAGIS in diesem Band und SCHUDNAGIS & WOMSER-HACKER 2002) sollen Erkenntnisse darüber gewonnen werden, wie man Lernsysteme besser benutzbar machen kann. Dabei werden System-Prototypen mit unterschiedlicher lerntheoretischer Basis konzipiert und hinsichtlich ihrer Bedienfreundlichkeit (aber auch im Hinblick auf erzielte Lerneffekte) mit Hilfe von Benutzertests miteinander verglichen. Daraus sollen Prinzipien für die Gestaltung von Lernsystemen für einen bestimmten Lernbereich (hier: Übung zur Vorlesung) abgeleitet werden. Der folgende Abschnitt skizziert die wesentlichen Elemente zweier System-Prototypen, die Aspekte verschiedener Lerntheorien miteinander verbinden.

2 Gegenüberstellung der Systemkonzepte

Das System *bekog* orientiert sich sowohl an der behavioristischen als auch an der kognitivistischen Auffassung von Lernen, während das System *kogkons* neben kognitivistischen auch konstruktivistische Elemente integriert. Die Systeme lassen sich im Hinblick auf folgende Punkte charakterisieren:

Systemstruktur: *bekog* ist weitgehend linear strukturiert; Aufgaben folgen auf die Präsentation kurzer Texte. Dagegen besteht *kogkons* aus zwei eigenständigen Bereichen (Themen- und Arbeitsbereich), die auch getrennt zu bearbeiten sind. Beide Oberflächen setzen Farben ein, um Orientierungshilfe in Form einer Farbleitlinie zu geben, gleichzeitig aber auch um eine positive Grundstimmung zu erzeugen.

Navigation: In *bekog* besteht neben einem direkten Zugriff über das Inhalts-

¹ Gefördert von der VW-Stiftung des Landes Niedersachsen unter CeBu-Hi.



verzeichnis auch die Möglichkeit, mit Hilfe von Pfeiltasten seitenweise vor- und zurück bzw. an den Anfang oder das Ende zu blättern. Daneben stehen eine Reihe nicht-linearer Links zur Verfügung (z.B. ins Glossar). In *kogkons* existieren zusätzlich noch inhaltlich motivierte Verknüpfungen zwischen Themen- und Aufgabenbereich, sowie eine weitergehende Strukturierung des Arbeitsbereichs, die an der Oberfläche mit Hilfe der Karteikarten-Metapher verdeutlicht wird. Dadurch wird jede Seite innerhalb des Arbeitsbereichs direkt zugänglich. Da aufgrund dieser vielfältigen Navigationsmöglichkeiten (zumindest anfangs) die Gefahr der Überforderung der Lernenden besteht, wird zusätzlich ein ‚idealer Pfad‘ angeboten, der den Weg durch *kogkons* weist.

Aufmerksamkeitssteuerung: Um die Aufmerksamkeit der Lernenden zu gewinnen, ist das System *bekog* mit kognitionspsychologisch motivierten Elementen (advance organizer, Begriffszuordnungsspiel, etc.) angereichert, die vorhandenes Wissen aktivieren sollen. Dagegen verwendet *kogkons* zur Steuerung der Aufmerksamkeit eine authentische, komplexe Aufgabenstellung, um einen Anreiz zur Auseinandersetzung mit dem Thema zu geben.

Reihenfolge Thema-Aufgabe: Während *bekog* durch eine strikte Reihenfolge von Texten und Aufgaben geprägt ist, bei der zunächst Wissen vermittelt und dieses dann angewendet wird, sind in *kogkons* Text und Aufgabe gleichberechtigt nebeneinander gestellt. Verfolgt man jedoch den Pfad, werden oftmals Aufgaben vor dem erklärenden Text präsentiert. Auf diese Weise sollen die Lernenden auf Probleme stoßen, diese erkennen und dafür Lösungen erarbeiten. Der Text dient im Anschluss daran als Feedback.

Aufgabenlösung (Hilfe und Feedback): Die Hilfe ist bei beiden Systemen gleich gestaltet, das Feedback variiert jedoch. Während in *bekog* der Benutzer die richtige Lösung anfordert, gibt *kogkons* automatisch Hinweise, dass ein Fehler aufgetreten ist, ohne stets gleichzeitig eine Lösung vorzuschlagen.

3 Ausblick

Mit Hilfe eines Benutzertests wurden die Stärken und Schwächen der Systeme, Benutzerpräferenzen, etc. ermittelt. Die Ergebnisse dienen als Basis für eine System-Überarbeitung als nächste Stufe unserer iterativen Vorgehensweise, die letztendlich in einer Zusammenführung der beiden Systemkonzepte münden soll.

4 Literatur

Schudnagis, M.; Womser-Hacker, Ch. (2002). Multimediale Lernsysteme softwareergonomisch gestalten: das Projekt SELIM. Tagungsband Mensch & Computer 2002.



ConceptManager: Pflege multilingualer Ontologien im crosslingualen Retrieval

A. Jackson, M. Lewandowski, Gr. Thurmair, J. Zwickl

Comprendium GmbH

Balanstr. 54

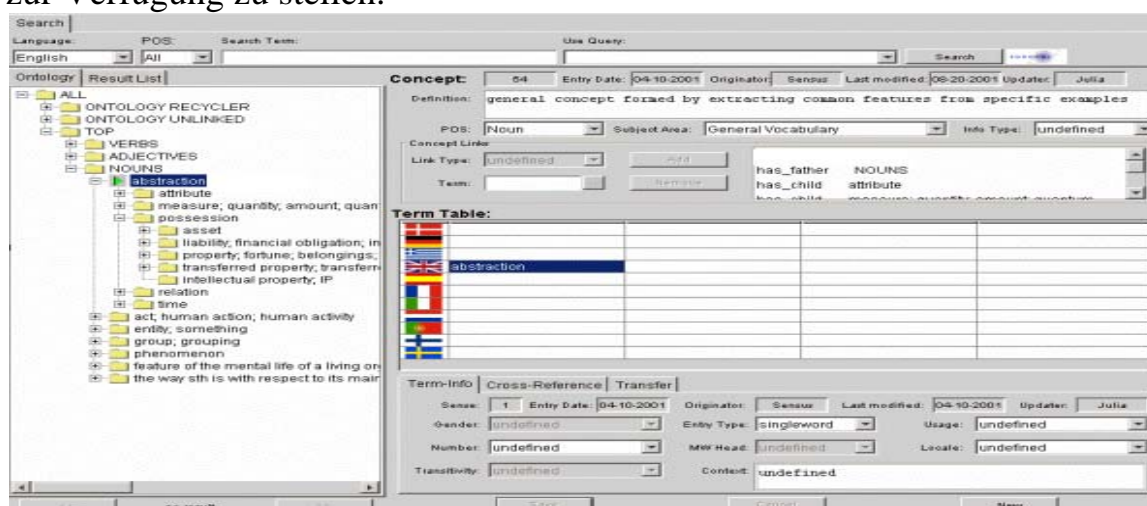
81541 München

Zusammenfassung

Der ConceptManager ist ein Tool zur Pflege multilingualer Ontologien, wie sie beim crosslingualen Retrieval benötigt werden. Er unterstützt die Erstellung von Begriffshierarchien und das Definieren von Begriffen; er unterstützt die Zuordnung von Termen zu diesen Begriffen in verschiedenen Sprachen; und er unterstützt die Zuordnung von linguistischen Annotationen zu diesen Termen, wie sie für die maschinelle Verarbeitung relevant sind.

1 Ontologien

Ontologien haben sich als nützliches Mittel bei der Suche entwickelt, weil sie eine Expansion der Suchbegriffe gestatten, die die Ergebnisse verbessern hilft. Trotz einiger Ansätze zur automatischen Erzeugung von Ontologien aus Korpora ist es in der Praxis erforderlich, eine interaktive Pflegekomponente zur Verfügung zu stellen.



Existierende Tools zur Pflege von Ontologien unterstützen üblicherweise den hierarchischen Aspekt, jedoch nicht den linguistischen Aspekt: Während Thesauri eine Normierung der Begrifflichkeit anstreben, muss im Retrieval die



Vielfalt der sprachlichen Realisierungen eines Begriffs erfasst und repräsentiert werden.

Der ConceptManager unterstützt die Erstellung von Hierarchiebäumen (*is_a*-Relationen) und die Verknüpfung der Begriffe mittels der bei EuroWordNet definierten Relationen. Die Begriffe, die die einzelnen Knoten bilden, werden durch Definition, Wortart und Fachgebiet näher beschrieben.

2 Terme, Multilingualität

Jeder Begriff kann durch Terme in verschiedenen Sprachen ausgedrückt werden. Der ConceptManager gestattet die Eingabe von Termen und ihrer Synonyme in den verschiedenen Sprachen; auf diese Weise entsteht ein multilinguales ConceptNet. Die Strategie dabei ist, eine universelle Ontologie mit mehrsprachigen Termen anzureichern, statt mehrere einzelsprachliche Ontologien einander zuzuordnen. Homographe Terme werden über ihre Beziehung zu den Begriffen disambiguiert.

3 Linguistische Beschreibung

Jeder Term einer Sprache kann um linguistische Annotationen angereichert werden: Wortart, Genus, Transitivität usw.; Querverweise wie „*ist_Abkürzung von*“; Übersetzungshinweise wie „*präferierte Übersetzung, Teiläquivalenz*“ können angegeben werden. Ziel ist es, das multilinguale ConceptNet möglichst vielen linguistischen Komponenten, von Indexing über Information Extraction bis zur maschinellen Übersetzung, zugänglich zu machen. Die linguistischen Features stellen ein Subset von *OLIF* dar, einem offenen Lexikonstandard, den der ConceptManager als Schnittstelle zu Fremdanwendungen bei Import und Export unterstützt.

4 Einbettung

Der ConceptManager ist Teil eines Workflows, in dem mittels Term-Extraktionsprogrammen aus einem repräsentativen Korpus die relevanten Suchterme gefunden werden. Sie werden in den ConceptManager importiert, dort interaktiv in ihre begrifflichen Relationen gebracht, ggf. um Übersetzungen angereichert und mit linguistischen Annotationen versehen. Das Ergebnis wird in die Datenstrukturen der weiterverarbeitenden Komponenten (z.B. einer Suchfragen-Expansions-Komponente) kompiliert. Das Tool ist erfolgreich zum Aufbau eines elfsprachigen ConceptNets mit ca. 15000 Begriffsknoten eingesetzt worden.



Implementierung und Evaluierung des MIMOR-Modells: Mehrfachindexierung zur dynamischen Methoden-Objekt-Relationierung im Information Retrieval

Ralph Kölle, Thomas Mandl, Christa Womser-Hacker

Informationswissenschaft - Universität Hildesheim
Marienburger Platz 22 - 31141 Hildesheim
{koelle, mandl, womser}@uni-hildesheim.de

1 Das MIMOR Modell

Das MIMOR-Modell (Mehrfachindexierung zur dynamischen Methoden-Objekt-Relationierung im Information Retrieval) ist ein Ansatz zur Erhöhung der Adaptivität auf einer Meta-Ebene. MIMOR arbeitet mit Mehrfachindexierung, also der Repräsentation der Objekte (Texte, Bilder usw.) aus mehreren Perspektiven. Große Evaluierungs-Studien haben u.a. gezeigt, dass die besten IR-Verfahren sich in der Qualität nur unwesentlich unterscheiden. Ihre Treffermengen weisen jedoch nur wenig Überschneidung auf. Das bedeutet, dass gute Verfahren etwa gleich viele relevante Dokumente finden, aber eben verschiedene (cf. WOMSER-HACKER 1997). Sogenannte Mehrfachindexierungs- und Fusionsansätzen versuchen dies auszunutzen und die relevanten Dokumente mehrerer Ansätze zu vereinigen. Das Problem besteht darin, die relevanten Dokumente in der Gesamtergebnismenge zu finden und stärker zu gewichten. MIMOR lernt, die Fusion durch Relevanz-Feedback und Benutzerpartizipation zu steuern. Das System kombiniert die Ergebnisse der einzelnen Systeme im Lauf des Lernprozesses zunehmend besser. Jedes System trägt mit einem bestimmten Gewicht zum Gesamtergebnis bei. Dieses Gewicht erhöht sich, wenn ein einzelnes System häufig positiv bewertete Dokumente auf die gesamte Ergebnisliste bringt (cf. MANDL & WOMSER-HACKER 2000, WOMSER-HACKER 1997).

2 Evaluierung mit CLEF Daten

Das MIMOR-Modell wurde in JAVA implementiert und konnte somit evaluiert werden. Für die Evaluierung von Information Retrieval Systemen haben sich die großen Evaluierungsinitiativen etabliert. In Europa ist dies vor allem die CLEF Kampagne (Cross Language Evaluation Forum, cf. KLUCK ET AL.



2002). Innerhalb der CLEF Daten sind insbesondere die GIRT Daten (German Indexing and Retrieval Testdatabase) interessant. Dabei handelt es sich um ca. 80.000 sozialwissenschaftliche Fachtexte, die auch intellektuell erschlossen sind.

MIMOR wurde mit den GIRT Daten evaluiert (cf. HACKL ET AL. 2002). Für einen Fusionsansatz müssen zunächst mehrere einzelne Retrievalsysteme zur Verfügung stehen, deren Ergebnisse verschmolzen werden. Die Wahl fiel auf das *irf* Paket (Information Retrieval Framework) des NIST (National Institute of Standards and Technology, USA), da dieses ebenfalls in JAVA realisiert ist und als Quellcode vorliegt. Mit *irf* konnten zwei Systeme simuliert werden, da es zahlreiche Parametrisierungsmöglichkeiten bietet. Diese wurden einmal mit gleichen Gewichten und einmal mit optimierten Gewichten linear fusioniert. Die Optimierung basiert auf einer Analyse des gleichen Systems mit den Daten und den menschlichen Relevanzurteilen des Jahres 2001. Nach Bekanntgabe der Ergebnisse durch die CLEF Organisatoren lässt sich der Einfluss der Fusion und der Optimierung abschätzen. Im nächsten Jahr sind die Einbindung mehrerer Systeme und die Teilnahme am mehrsprachigen Retrieval geplant.

Die MIMOR Implementierung eignet sich aufgrund der Flexibilität des objekt-orientierten Ansatzes auch für die weitere Untersuchung der Validität der Testergebnisse von CLEF.

3 Literaturverzeichnis

- HACKL, René; KÖLLE, Ralph; MANDL, Thomas; WOMSER-HACKER, Christa (2002): Domain Specific Retrieval Experiments at the University of Hildesheim with the MIMOR System. Erscheint in: Cross Language Evaluation Forum. CLEF-Workshop Notes. 19.-20.9.2002. Rom.
- KLUCK, Michael; MANDL, Thomas; WOMSER-HACKER, Christa (2002): Cross-Language Evaluation Forum (CLEF): Europäische Initiative zur Bewertung sprachübergreifender Retrievalverfahren. In: nfd Information - Wissenschaft u. Praxis vol. 53 (2). S. 82-89.
- MANDL, Thomas; WOMSER-HACKER, Christa (2000): Ein adaptives Information Retrieval Modell für Digitale Bibliotheken. In: Knorz, Gerhard; Kuhlen, Rainer (Hrsg.): Informationskompetenz – Basiskompetenz in der Informationsgesellschaft. Proc 7. Intl Symposium für Informationswissenschaft (ISI 2000) 8.-10. Nov., Darmstadt. S. 1-16.
- WOMSER-HACKER, Christa (1997): Das MIMOR-Modell. Mehrfachindexierung zur dynamischen Methoden-Objekt-Relationierung im Information Retrieval. Habilitationsschrift. Universität Regensburg, Informationswissenschaft. 1997.



MURBANDY - Benutzerfreundliche Recherche von Landnutzungsdaten

Bernd Hermes, Maximilian Stempfhuber

Informationszentrum Sozialwissenschaften
Lennéstr. 30
D-53113 Bonn
stempfhuber@bonn.iz-soz.de

Abstract

Landnutzungsdaten werden – wie viele andere Faktendaten auch – vorwiegend über inhaltliche und thematische Aspekte gesucht und nicht über den bloßen Abgleich von Suchparametern mit Attributen von Datensätzen. Die thematischen Aspekte zur Beschreibung der Datensätze werden Nomenklaturen und Klassifikationen entnommen und aufgrund ihrer Komplexität dem Benutzer zur Auswahl angeboten. MURBANDY verwendet in seiner Benutzungsoberfläche das Prinzip der dynamischen Anpassung und Visual Formalisms um die semantischen Beziehungen zwischen Nomenklaturen zu visualisieren und gleichzeitig für die detaillierte, iterative Suche verfügbar zu machen.

1 Einleitung

MURBANDY (Monitoring Urban Dynamics) ist ein Forschungsprojekt des Space Application Institute des Joint Research Centre der Europäischen Kommission (Ispra, Italien), bei dem Landnutzungsdaten von 25 europäischen Metropolen aus Luftbildern extrahiert und anhand der internationalen Landnutzungsklassifikation erschlossen wurden. Zur Nutzung dieser Daten über das Internet wurde am Informationszentrum Sozialwissenschaften (Bonn) eine Benutzungsoberfläche entworfen und in Java implementiert, die den Zugriff auf und die Evaluation dieser Daten unter einer nutzerfreundlichen grafischen Oberfläche erlaubt.

2 Oberflächendesign

Bei der Gestaltung der Benutzungsoberfläche wurden softwareergonomische Prinzipien wie dynamische Anpassung bzw. enge Kopplung von Bildschirm-



elementen und Query Previews kombiniert, um eine einheitliche, adaptive Oberfläche zu schaffen, die Anfänger wie Fortgeschritten ohne konzeptionelle Brüche möglichst optimal unterstützt. Dabei werden zur Anfrageformulierung Vorlagelisten verwendet, aus denen der Benutzer die gewünschten Landnutzungsarten und Städte auswählen kann. Die gewählten Begriffe werden in einer Zustandsanzeige gesammelt, so dass der Nutzer auf einen Blick erkennen kann, aus welchen Begriffen sich seine Anfrage zusammensetzt.

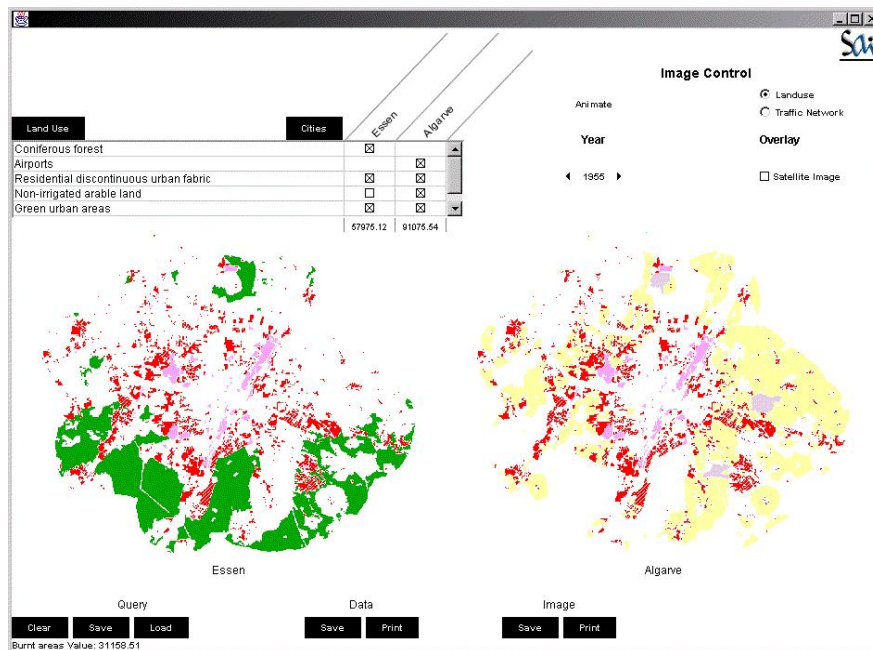


Abb. 1: MURBANDY Ergebnisdarstellung

Die Abbildung 1 zeigt den Ergebnisbildschirm, der eine Tabelle als Visual Formalism (Nardi & Zarmer 1993) enthält, bestehend aus den beiden Zustandsanzeigen für Landnutzungsarten und Städte. Der Visual Formalism stellt die semantischen Beziehungen zwischen beiden Suchattributen dar und ermöglicht gleichzeitig eine detaillierte Interaktion. Er wird damit selbst zum Eingabe und Explorationswerkzeug und ermöglicht die Bedienung als 1-Bildschirm-System.

3 Literatur

Nardi, B. A.; Zarmer, C. L. (1993). *Beyond Models and Metaphors: Visual Formalisms in User Interface Design*. In: *Journal of Visual Languages and Computing*, 1993, 4, pp. 5-33.



LINUCS - Eine Notation zur Unterstützung von Repräsentation und Retrieval spezieller chemischer Strukturen

Elke Lang¹, Andreas Bohne-Lang², Claus-Wilhelm von der Lieth², Alexander Loß²

¹Fachhochschule Darmstadt
FB Informations- und Wissensmanagement
Max-Planck-Str. 2
64807 Dieburg
lang@iuw.fh-darmstadt.de

²Deutsches Krebsforschungszentrum
Zentrale Spektroskopie
AG Molecular Modeling
Im Neuenheimer Feld 280
69120 Heidelberg
a.bohne@dkfz.de

Zusammenfassung

Molekülstrukturorientierte Suche ist eine der wichtigsten Anforderungen an Substanzinformationssysteme. Da die Suche inzwischen meist von Anwendungsfachleuten ohne Kenntnis spezieller Retrievalsprachen durchgeführt wird, ist eine komfortable, intuitiv verständliche Benutzeroberfläche ein wesentlicher Akzeptanzfaktor. Beim Substanzinformationssystem SWEET-DB bildet ein Strukturgraph-Parser das Bindeglied zwischen der graphisch orientierten Eingabemaske für Strukturvorgaben und der (linearen) LINUCS-Graphnotation, die als Speicherformat für Zuckermolekül-Strukturen dient.

Abstract

Structure-oriented retrieval is an essential feature of substance information systems. Chemists most often perform standard retrieval projects on their own and therefore prefer user surfaces that allow input according to chemical notation rather than requiring knowledge of special retrieval languages. SWEET-DB, a glycosubstance information system, offers structure retrieval by using an input matrix that is fully conform to IUPAC nomenclature. A structure parser converts graphical 2D structure input into linear notation that is used for storage and graph matching.

1 Die LINUCS-Notation im Kontext der SWEET-DB

Die SWEET-DB [Loß2002] ist ein Substanzinformationssystem, das Molekülstrukturen einer bestimmten Substanzklasse enthält (Kohlenhydrate). Diese Spezialisierung erlaubt den Gebrauch entsprechender Notationen, die effi-



zienter sind als substanzklassenunabhängige Beschreibungsverfahren. Derartige Spezialnotationen haben sich z.B. für die Deskription und Suche von Proteinsequenzen eingebürgert [Berman2000]. Seit einiger Zeit werden Versuche unternommen, ein derartiges Verfahren auch für Kohlenhydrate zu entwickeln, wobei deren höhere strukturelle Komplexität, vor allem die Verzweigung, erhebliche Schwierigkeiten verursacht [Laine1994]. Die meisten Verfahren verlangen über die Standard-Nomenklatur [IUPAC1997] hinaus die Einhaltung weiterer Konventionen [Engelsen1996]. Die LINUCS-Notation [Lieth2001] bietet die Möglichkeit, Suchstrukturvorgaben mit einer Eingabematrix zu erstellen, in die wie gewohnt zweidimensionale (verzweigte) Strukturgraphen eingetragen werden können, die aus Monomertyp- und Bindungstyp-Elementen kombiniert werden. Der LINUCS-Parser wird zum einen verwendet, um bei der Aufnahme neuer Substanzdatensätze deren lineare Strukturnotation als Speicherformat zu erzeugen. Zum anderen wandelt er bei der Struktursuche die graphisch erstellte Strukturvorgabe in die lineare Notation um. Diese wird anschließend zum Graphvergleich benutzt.

Die LINUCS-Notation ist durch ihre lineare Form rechnergeeignet und ermöglicht eine schnelle Struktursuche; auch Ähnlichkeits- und Teilstruktursuche sind möglich. Bei der Aufnahme neuer Daten bietet LINUCS die Möglichkeit, neu aufgenommene Strukturen auf Plausibilität zu untersuchen.

Das LINUCS-Verfahren und die SWEET-DB sind unter <http://www.dkfz-heidelberg.de/spec/> zugänglich, dort sind auch einführende Beispiele zu finden.

2 Literatur

- [Engelsen1996] Engelsen SB, Cors S, Mackie W, Pérez S: A Molecular Builder for Carbohydrates: Application to Polysaccharides and Complex Carbohydrates. *Biopolymers* 39 (1996) 417-433
- [IUPAC1997] IUPAC-IUBMB, Nomenclature of Carbohydrates. *Carbohydrate Research* 297 (1997) 1-92
- [Laine1994] Laine RA: Calculation of all possible oligosaccharide isomers both branched and linear yield 1.05×10^{12} structures of a reducing hexasaccharide. *Glycobiology* 4 (1994) 759-767
- [Lieth2001] von der Lieth CW, Bohne-Lang A, Lang E, Förster T: LINUCS: Linear Notation for Unique description of Carbohydrate Sequences. *Carbohydrate Research* 336 (2001) 1-11
- [Loß2002] Loß A, Bunsmann P, Bohne A, Loß A, Schwarzer E, Lang E, von der Lieth CW: SWEET-DB: an attempt to create annotated data collections for carbohydrates. *Nucl. Acids. Res.* 30 (2002) 405-408